



**HAL**  
open science

# Validity and Accuracy of Impulse-Response Models for Modeling and Predicting Training Effects on Performance of Swimmers

Thierry Busso, Sébastien Chalencon

► **To cite this version:**

Thierry Busso, Sébastien Chalencon. Validity and Accuracy of Impulse-Response Models for Modeling and Predicting Training Effects on Performance of Swimmers. *Medicine and Science in Sports and Exercise*, 2023, 55 (7), pp.1274-1285. 10.1249/mss.0000000000003139 . hal-04713855

**HAL Id: hal-04713855**

**<https://hal.science/hal-04713855v1>**

Submitted on 2 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**AMERICAN COLLEGE  
of SPORTS MEDICINE®**  
**LEADING THE WAY**

*. . . Published ahead of Print*

## **Validity and Accuracy of Impulse-Response Models for Modeling and Predicting Training Effects on Performance of Swimmers**

Thierry Busso<sup>1</sup> and Sébastien Chalencon<sup>2</sup>

<sup>1</sup>Université Jean Monnet Saint-Etienne, Lyon 1, Université Savoie Mont-Blanc, Laboratoire Interuniversitaire de Biologie de la Motricité, Saint-Etienne, FRANCE; <sup>2</sup>Club des Dauphins de Guilhaumand-Granges, Guilhaumand-Granges, FRANCE

Accepted for Publication: 29 January 2023

*Medicine & Science in Sports & Exercise*® **Published ahead of Print** contains articles in unedited manuscript form that have been peer reviewed and accepted for publication. This manuscript will undergo copyediting, page composition, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered that could affect the content.

# Validity and Accuracy of Impulse-Response Models for Modeling and Predicting Training Effects on Performance of Swimmers

Thierry Busso<sup>1</sup> and Sébastien Chalencon<sup>2</sup>

<sup>1</sup>Université Jean Monnet Saint-Etienne, Lyon 1, Université Savoie Mont-Blanc, Laboratoire Interuniversitaire de Biologie de la Motricité, Saint-Etienne, FRANCE; <sup>2</sup>Club des Dauphins de Guilhaud-Granges, Guilhaud-Granges, FRANCE

## Address for Correspondence:

Thierry Busso, Laboratoire Interuniversitaire de Biologie de la Motricité, Université Jean Monnet Saint-Etienne, LIBM - Campus Santé Innovations – IRMIS, 10, Rue de la Marandière, 42270 Saint-Priest-en-Jarez, France; Phone: +33 04 77 42 18 71; E-mail: busso@univ-st-etienne.fr

## Conflict of Interest and Funding Source:

The authors declare that they have no conflicting interests. The results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. The results of the present study do not constitute endorsement by the American College of Sports Medicine.

## ABSTRACT

**Purpose:** The aim of this study was to compare the suitability of models for practical applications in training planning.

**Methods:** We tested 6 impulse-response models including Banister's model (Model Ba), a variable-dose response model (Model Bu) and indirect-response models differing in the way they account or not for the impact of prior training on the ability to respond effectively to a given session. Data from 11 swimmers were collected during 61 weeks across two competitive seasons. Daily training load was calculated from the number of pool-kilometers and dry land workout equivalents, weighted according to intensity. Performance was determined from 50-m trials done during training sessions twice a week. Models were ranked on the base of Aikake's information criterion along with measures of goodness-of-fit.

**Results:** Models Ba and Bu gave the greatest Akaike weights,  $0.339 \pm 0.254$  and  $0.360 \pm 0.296$  respectively. Their estimates were used to determine the evolution of performance over time after a training session and the optimal characteristics of taper. The data of the first 20 weeks were used to train these two models and predict performance for the following 8 weeks (validation dataset 1) and for the following season (validation dataset 2). The mean absolute percentage error between real and predicted performance using Model Ba was  $2.02 \pm 0.65$  and  $2.69 \pm 1.23$  % for validation dataset 1 and 2 respectively and  $2.17 \pm 0.65$  and  $2.56 \pm 0.79$  % with Model Bu.

**Conclusions:** The findings showed that, although the two top-ranked models gave relevant approximations of the relationship between training and performance, their ability to predict future performance from past data was not satisfactory for individual training planning.

**Key Words:** MODEL SELECTION, FATIGUE, TRAINING ADAPTATION, TAPERING

ACCEPTED

## INTRODUCTION

There is renewed interest in mathematical models of training effects on performance to guide training planning (1-5). The most widely used models are impulse-response models which were especially suited to acquiring knowledge on taper for performance peaking (6). However, progress towards implementing models in new applications for designing or monitoring individual athlete's training requires further exploration of model accuracy in predicting performance.

The use of impulse-response models in athletics was initiated by Banister and co-workers in order to characterize the temporal effects of training on performance (7, 8). Based on the mathematical formulation of the so-called impulse response, this model characterizes temporal variations in performance following a single training bout. For a given training intervention, the change in performance at a given day is obtained by summing over time the response to each single bout done the days before. Calculating performance for a period of intervention requires estimating model parameters by fitting the model to data on real performances. The data sets used to solve the model must include, on one hand, daily quantification of the amount of training, generally referred to as training loads (TL). These loads are expressed in an arbitrary unit (training unit in this study) depending on the metrics used to aggregate the different forms of exercise making up the training sessions. On the other hand, the data must also provide results of performance tests done regularly during the intervention so that the model can be parametrized and we can test its ability to describe the relationship between training and performance. It has been shown that measurement reliability and testing frequency (9) and methods for parameter

estimation (10, 11) can affect model accuracy. The number of athletes providing the data and the duration of data collection are also key to producing a powerful analysis.

Various impulse-response models have attempted to capture the main features of the relationship between variations of training effort and temporal variations of performance (7, 8, 12-14). In the original model of Banister et al (7), Model Ba in this study, the impulse-response of performance is the result of body adaptation (also referred to as fitness) and fatigue which act positively and negatively on performance respectively. Both positive and negative effects are modelled in an identical fashion using first-order kinetics (Fig. 1 panels A and B). It results that performance decreases following training, requiring time, referred to as  $t_n$  to recover to its initial level before a peak  $p_g$  is attained at the time  $t_g$  after which performance enhancement dissipates.

The model referred to as Model Bu in this study is an extension of Model Ba. It considers that the capacity to benefit from training is impaired by training overload and can be restored by reducing training (12). Contrary to Model Ba, Model Bu is based on the idea that the effects of a given training bout are dependent on prior training efforts. More precisely, the negative effect of a single session, *i.e.* an isolated bout, is lower than the effect of repeated similar bouts, *i.e.* a last bout effect. This brings on changes in  $t_n$  and  $t_g$  according to prior training (Fig. 1 panels C and D). Model Bu outperformed Model Ba using data sets from volunteers in a training program designed for this purpose (12). However, when using data from elite athletes, the indicators of goodness-of-fit were in favor of Model Ba (15, 16).

In Model Ba and Model Bu, the gain in adaptation is entirely produced immediately after

the training session and dissipates with a first-order rate constant. It is recognized, however, that adaptation results from physiological processes intervening during post-exercise recovery. A better description of the delay in adaptation following training sessions is obtained by using double exponential to obtain a positive effect (8, 14, 15) or indirect response model (13). The later was inspired by pharmacodynamics, considering the positive effect to be the result of the transformation of a signal secondary to the primary training stimulus (Fig. 1 panel E). With indirect response models, negative effects of training were better described by an inhibition of production of positive effects due to prior training sessions (Model TI, Fig. 1 panel F) rather than fatigue with a first-order kinetics with or without inhibition process, Model TIF and Model TF respectively). Nevertheless, it is likely that the absence of a fatigue component in the model makes it unsuitable for athletes whose performance capacity diminishes when their training is intensified. Furthermore, the inclusion of a fatigue factor, which changes kinetics according to prior training (Model TF2) was not tested in combination with an inhibition of the adaptation process.

Selecting a model as the best approximation among a set of models is a basis for statistical inference because it tells us what effects represented by which parameters are best supported by the data (17). Models differ in complexity *i.e.* in the number of parameters to be identified by fitting the model to data. An under-fitted model may not adequately capture the relationship between temporal variations in training and performance. Conversely, an over-fitted model will tend to reproduce irrelevant variance in the training data, increasing the variability of estimates. The best model is ideally a parsimonious model able to capture the true relationship between the variables of interest while not over-fitting the data. Often  $R^2$  adjusted and F-ratio



tests are used to verify if an increase in model complexity is statistically relevant and would not lead to overfitting (12, 15, 16, 18). However Akaike's information criterion (AIC) is recognized as being a more powerful way to identify the best model because it ranks the models from best to worst by weighting the evidence that each model is likely the best (17, 19).

Another way to test models is to measure their ability to predict future performance from past data used to parametrize or "train" the model. Both Model Ba and Model Bu were found to produce good predictions of performance in swimmers (20). This study used the data from a 15-wk training cycle for train the model (training data set) and those of the following 15-wk training cycle for testing the prediction (validation data set). The proximity in time between the 2 training cycles does question, however to what point conclusions drawn for these observations can be generalized.

Our first aim was to generate large data sets from a group of swimmers with reliable and frequent performance measurements during an extended period to have substantial statistical power for estimating model parameters and testing the ability of the models to predict performance. Secondly, we aimed to rank a set of impulse-response models using AIC and their extensions along with indicators of goodness-of fit. The estimates of the selected models were then used to explore the temporal variation of performance following a training session and the characteristics of the optimal taper period for comparison with data in the literature. Finally, the data from each athlete were split into training and validation data sets to determine the ability of the models to predict future performance.

## METHODS

### Data collection and processing

Data came from 11 swimmers including 6 females specialized in 50-m or 50 and 100-m events (8 in freestyle and 3 in other strokes). They were aged  $17.3 \pm 2.9$  years with competitive experience of  $7.1 \pm 2.9$  years at the beginning of this study. Best performances achieved during 50-m trials in 25-m pool were  $75.9 \pm 4.3$  % of world-record speeds. All swimmers provided informed, written consent prior to data use in accordance with procedures approved by Institutional Review Board of University Hospital of Saint-Etienne (reference number IRBM1262021/CHUSTE).

A single coach prescribed the training program for the entire group of swimmers and collected the data related to training efforts and performance testing from January to July 2019 (Season 1) and, then, for 8 of them until March 2020 when training ceased due to Covid 19 lockdown (Season 2). Season 1 covered the summer season comprising 3 training cycles ranging from weeks 1 to 28 before training was suspended for summer break. Season 2 covered the following winter season from weeks 34 to 48 (training cycle 4), then a break at Christmas before the beginning of the summer season from weeks 52 to 61 (training cycle 5).

Swimming speed during a 1500-m time trial ( $S_{15}$ ) was used to establish individual timetables of time per distance according to intensity levels : level 1 for active recovery (<90% of  $S_{15}$ ), level 2 for basic endurance training (90-94% of  $S_{15}$ ), level 3 for aerobic training (95-99% of  $S_{15}$ ), level 4 for lactate threshold (100-104% of  $S_{15}$ ), level 5 for maximal oxygen uptake (105-109% of  $S_{15}$ ), level 6 for anaerobic capacity (110-120% of  $S_{15}$ ) and level 7 for sprints at maximal

speed. Conditioning training comprised dry land exercises. Stroke rate and resistance settings were determined by the coach in correspondence with swimming speed, allowing him to design the training sessions according to intensity level. Training workouts were converted into equivalent swim distance at each intensity level.

Training load (TL) was quantified each day from the weighted sum of the number of pool-kilometers swum and the dry land workout equivalent at each intensity level. Distances swum at zone 1 intensity were not included in the computation. The weighting factors for the remaining intensity zones were those proposed by Mujika et al (21) as follows:

$$TL = Z_2 + 2 \cdot Z_3 + 3 \cdot Z_4 + 5 \cdot Z_5 + 8 \cdot (Z_6 + Z_7) \quad (1)$$

where  $Z_2$ ,  $Z_3$ ,  $Z_4$ ,  $Z_5$ ,  $Z_6$  and  $Z_7$  are the number of kilometers swum at intensity levels 2, 3, 4, 5, 6 and 7 respectively. TL was set to 0 during periods of Christmas and summer breaks.

Swimmers trained every day from Monday to Friday and rested the week-end except for participating in competitions participation. Training sessions of each Monday and Friday included a 50-m time trial in swimmer's stroke specialty in a 25-m pool. Each trial was supervised by a coach who blew a first whistle to position the swimmer on the starting block, a second whistle to give the start signal, and then manually timed the 50-m trial with a stopwatch (Seiko, Japan). Performance was assessed from the average speed during the trial expressed as a

percentage of the first value of Season 1. The reliability of this test assessed from the individual coefficient of variation from the first four trials of this study was  $0.40 \pm 0.23$  %.

### Model computations

Six models were solved for each swimmer from daily TL and performance measured twice a week (see Supplemental Figure, Supplemental Digital Content 1, which gives a schematic representation of each model, <http://links.lww.com/MSS/C802>).

TL is considered as a discrete function, *i.e.*, a series of impulses each day,  $w^i$  on day  $i$ , and the model performance on day  $i$  was estimated by mathematical recursion from the series of  $w$  before day  $i$ . The equations of each tested model were written as recursive sequences in which each term on a given day was defined as a function of the terms on either the same or the preceding day.

The models other than Model TI assume that performance is the result of change from baseline ( $p_{base}$ ) because of cumulated adaptation to training (positive effect :  $pe$ ) and cumulated fatigue (negative effect :  $ne$ ).

$$\hat{p}^n = p_{base} + pe^n - ne^n \quad (2)$$

where  $pe^n$ ,  $ne^n$  and  $\hat{p}^n$  are the estimations of  $pe$ ,  $ne$  and performance on day  $n$  respectively with  $pe^0$ ,  $ne^0$  initialized to 0.

Model Ba initiated by Banister et al is based on the idea that each training bout contributes to an increase in positive and negative effects proportional to corresponding  $w$  with multiplying factors ( $kin_{pe}$  and  $kin_{ne}$  respectively  $pe$  and  $ne$ ) which decay away in the following days in an exponential fashion at separate rate constants ( $kout_{pe}$  and  $kout_{ne}$  respectively). Model Ba is thus defined by the following equations

$$pe^n = \sum_{i=1}^{n-1} (kin_{pe} \cdot w^i + pe^{i-1} \cdot e^{-kout_{pe}}) \quad (3)$$

$$ne^n = \sum_{i=1}^{n-1} (kin_{ne} \cdot w^i + ne^{i-1} \cdot e^{-kout_{ne}}) \quad (4)$$

Model Bu assumes that the multiplying factor for fatigue  $kin_{ne}$  is itself considered to increase by a quantity proportional to each  $w$  and decay exponentially away. The value of  $kin_{ne}$  at day  $i$  is estimated by mathematical recursion using a first-order filter with a gain terms  $kin_{in\_ne}$  and a rate constant  $kout_{in\_pe}$  as follows

$$kin_{ne}^i = \sum_{j=1}^i (kin_{in\_ne} \cdot w^j + kin_{ne}^{j-1} \cdot e^{-kout_{in\_ne}}) \quad (5)$$

with  $kin_{in\_ne}^0$  initialized to 0.

Cumulated negative effect at day  $n$  is computed with Model Bu as follows:

$$ne^n = \sum_{i=1}^{n-1} (kin_{ne}^i \cdot w^i + ne^{i-1} \cdot e^{-kout_{ne}}) \quad (6)$$

Exponential decay of training effect is characterized by rate constant rather than time constant designated by  $\lambda$  which is its reciprocal.

The other models tested in this study are based on an indirect response because  $pe$  result from a secondary signal produced by primary training stimulus. The secondary signal  $s$  accumulates with training and fades away exponentially as follows:

$$s^n = \sum_{i=1}^n (kin_s \cdot w^i + s^{i-1} \cdot e^{-kout_s}) \quad (7)$$

where  $kin_s$  is the multiplying factor for the secondary signal and  $kout_s$  is the rate constant with  $s^0$  initialized to 0.

In Model TF and Model TF2,  $pe$  at a given day is assumed to increase by a quantity proportional to  $s$  the day before and decay with a first-order dynamic as follows:

$$pe^n = s^{n-1} + pe^{n-1} \cdot e^{-kout_{pe}} \quad (8)$$

In Model TI and Model TIF, training acts negatively by inhibiting the secondary signal that drives the positive effect. The variable  $Inhib$  was added to describe this inhibition so that its value at day  $n$  is proportional to  $w^n$  as follows:

$$Inhib^n = k_{inhib} \cdot w^n \quad (9)$$

where  $k_{inhib}$  is the multiplying factor for inhibition process.

The production of  $pe$  varies in function of  $Inhib$  resulting in the modification of Eq.8 as follows:

$$pe^n = s^{n-1} \cdot (1 - Inhib^{n-1}) + pe^{n-1} \cdot e^{-kout_{pe}} \quad (10)$$

Performance with Model TI is  $p_{base}$  added to  $pe$  estimated by Eq. 9 *i.e.* without fatigue. In Model TF and Model TF2,  $pe$  is estimated from Eq. 7, *i.e.* without an inhibition process, but counterbalanced by fatigue estimated from  $ne$  given by Eq. 4 as in Model Ba (Model TF) or by Eq. 5 and 6 as in Model Bu (Model TF2). Finally, performance with Model TIF is estimated from the balance between  $pe$  with inhibition process given by Eq. 7 to 10) and  $ne$  given by Eq. 4.

### Model fitting

The model parameters were determined by minimizing the residual sum of square (RSS) between estimated and measured performance as follows:

$$RSS = \sum_{i=1}^n (p^i - \hat{p}^i)^2 \quad (11)$$

with  $n$  the number of performance tests used to fit the model. RSS was minimized by using a hybrid method with combines standard linear regression with an L-BFGS-B algorithm which is a quasi-Newton method allowing us to set lower and upper bounds on the estimates. The latter was implemented using `constrOptim` function in R package (22) for fitting the rate constants ( $kout_{pe}$ ,  $kout_{ne}$ ,  $kout_{in\_ne}$  and  $kout_s$  according to the model) and  $k_{inhib}$  for Model TI and TIF. At each step of the optimization, the other parameters (*i.e.* the multiplying factor  $kin_{pe}$ ,  $kin_{ne}$  and  $kin_{in\_ne}$  according to the model) were obtained by linear regression directly. This hybrid method allowed us to define the optimization starting point of and the search region for only a part of the entire set of model parameters. We verified that the solution given by this optimization method

were not sensitive to the starting point of the computation. The details of the minimization procedure can be found in the supplementary information (see document, Supplemental Digital Content 2, which gives R script of the models, <http://links.lww.com/MSS/C803>).

To rank the models according to goodness-of-fit and AIC results, they were fitted to all the data collected for each swimmer giving  $n=56$  for 3 swimmers (Season 1 only) and  $n=104$  for 8 swimmers (Seasons 1 and 2). For testing their ability to predict performance from past data, models were fitted to the data of training cycles # 1 and 2 (training data sets with  $n=40$  for each swimmer). The estimates for Model Ba and Model Bu were used to predict performance during cycle # 3 for 11 swimmers and during cycles # 4 and 5 for 8 swimmers.

### Model estimates

In addition to model parameters, variables were calculated to characterize the time response of performance to a single training bout.  $t_n$ , the time to recover performance and  $t_g$ , the time to peak performance after training completion were computed from parameters of Model Ba as follows:

$$t_n = \frac{\ln(\text{kin}_{ne}/\text{kin}_{pe})}{\text{kout}_{ne} - \text{kout}_{pe}} \quad (12)$$

and

$$t_g = \frac{\ln(\text{kin}_{ne} \cdot \text{kout}_{ne} / \text{kin}_{pe} \cdot \text{kout}_{pe})}{\text{kout}_{ne} - \text{kout}_{pe}} \quad (13)$$

$p_g$  the maximal gain in performance after a training session at a given TL is given by:



$$p_g = \text{kin}_{pe} \cdot TL \cdot e^{-k_{out_{pe}} \cdot t_g} - \text{kin}_{ne} \cdot TL \cdot e^{-k_{out_{ne}} \cdot t_g} \quad (14)$$

Calculation of these variables for Model Bu must account for variations in  $\text{kin}_{ne}$  according to cumulated training. Its lowest value is that for a single training bout dependent on corresponding TL (single session). If the same training session is repeated every day,  $\text{kin}_{ne}$  will tend towards a limit value dependent on both TL and  $k_{out_{in_e}}$  (asymptote). Both estimates were estimated as follows:

$$\text{kin}_{ne}(\text{single session}) = \text{kin}_{ne} \cdot TL \quad (15)$$

and

$$\text{kin}_{ne}(\text{asymptote}) = \frac{\text{kin}_{ne} \cdot TL}{1 - e^{-k_{out_{in_{ne}}}}} \quad (16)$$

## Statistics

Let  $p$  be the number of parameters of each model: 5 for Model Ba and TI, 6 for Model Bu and TF, 7 for Model TF2 and TIF. We calculated the following measures of goodness-of-fit(19):

- the coefficient of determination ( $R^2$ ):

$$R^2 = 1 - \frac{RSS}{TSS} \quad (17)$$

with  $TSS$  the total sum of squares. The mean square error (SE) was computed as the square root of  $RSS/(n-p)$ . F-ratio was computed to assess the statistical significance of one model fit with

degrees of freedom (DF) set to  $p-1$  for the model and  $n-p-1$  for the residuals.

- the mean absolute percentage error (MAPE)

-

$$MAPE = \sum_{i=1}^n \frac{|p^i - \hat{p}^i|}{\hat{p}^i} \cdot 100$$

- the adjusted coefficient of determination ( $Adj.R^2$ ):

-

$$Adj.R^2 = 1 - \frac{n-1}{n-p} \cdot (1 - R^2) \quad (18)$$

- the bias-corrected Akaike Information Criterion ( $AICc$ ):

-

$$AICc = 2 \cdot p - 2 \cdot \ln(L) + \frac{2p(p+1)}{n-p-1} \quad (19)$$

with  $\ln(L)$  the log-likelihood of the estimated model computed as follows:

$$\ln(L) = -0.5 \cdot n \cdot \left[ \ln(2\pi) + 1 - \ln(n) + \ln \left( \sum_{i=1}^n (p^i - \hat{p}^i)^2 \right) \right] \quad (20)$$

- the Akaike weights,  $w_i(AICc)$  provide weight-of-evidence measurements for each of the six models tested in this study:

-

$$w_i(AICc) = \frac{\exp(-0.5 \cdot \Delta_i(AICc))}{\sum_{k=1}^6 \exp(-0.5 \cdot \Delta_k(AICc))} \quad (21)$$

where  $i$  is the model number and  $\Delta_i(AICc)$  the difference between  $AICc$  of model  $i$  and the lowest  $AICc$ .

One-way repeated measures analysis of variance (ANOVA) was conducted to compare between models of measures of goodness-of-fit and characteristics of impulse response. The reliability of the prediction of performances for the testing data sets was assessed from the calculation of systematic error, typical error and MAPE between predicted and measured performances. Two-way mixed-effect ANOVA was done to compare between models and data sets. All analyses included Greenhouse–Geisser correction to ensure sphericity assumption. When significant differences were found, Tukey’s post-hoc test was used for pairwise comparisons. Paired  $t$  test was conducted for other comparisons when appropriate. The significance level was set at  $P < 0.05$ .

## RESULTS

### Data sets

Figure 2 panels A and B give weekly TL for Season 1 and Season 2 respectively. Weekly TL were higher during the first part of each training cycle with a reduction during the 2 or 3 last weeks. The performance was maintained or decreased during the first part of each training cycle and then increased when TL was reduced until a peak before the beginning of the following cycle (Fig. 2 panels C and D for Season 1 and 2 respectively).

## Model fits

Table 1 summarizes the indicators of goodness-of-fit of each model fitted to the data of the 2 seasons of training in each swimmer.  $R^2$ ,  $\text{adj-}R^2$  and SE from Model TI was significantly lower than from other tested models ( $P < 0.05$ ). F-ratio showed that Model TI only significantly fitted the performance of 7 swimmers ( $P < 0.05$  in 2 and  $P < 0.001$  in 5) whereas other models significantly fitted all of swimmers individually ( $P < 0.001$  for all data sets with the exception of 1 with Model TF and Model TIF which was at  $P < 0.05$ ).

No difference in  $R^2$ ,  $\text{adj-}R^2$  and SE was observed between the models other than for TI. We used  $w(\text{AIC})$  which includes model complexity (i.e. the number of model parameters) to rank the models. Individual  $w(\text{AIC})$  scores were best in 5 swimmers with Model Ba and in 6 swimmers with Model Bu. No difference was observed at the group level in  $w(\text{AIC})$  between Model Ba and Bu (Table 1). Based on these results, Model Ba and Model Bu rank equally as the best models among those tested in this study.

Figure 3 panel A shows the performance fit with Model Ba from the entire data set of one representative swimmer. It shows substantial overestimation of the lowest values and underestimation of peak values at the end of each training cycle. This tendency is observable for every swimmer using both Model Ba and Model Bu. At the group level, the difference between Model Ba and real performance was different from zero ( $P < 0.001$ ) at the end of each training cycle:  $-1.5 \pm 0.9$ ,  $-3.3 \pm 1.2$ ,  $-3.8 \pm 1.2$ ,  $-3.9 \pm 1.2$  and  $-4.1 \pm 1.3$  % at the end of cycle #1 to 5 respectively. Estimates for Model Bu also significantly underestimated peak performance ( $P < 0.001$ ):  $-1.9 \pm 1.2$ ,  $-3.2 \pm 1.4$ ,  $-4.0 \pm 1.3$ ,  $-3.8 \pm 1.2$  and  $-4.0 \pm 1.2$  % at the end of cycle #1 to

5 respectively. Residuals and real performance were highly correlated in each swimmer:  $0.84 \pm 0.06$  and  $0.84 \pm 0.04$  for Model Ba and Model Bu respectively.

Characteristics of performance over time after a training session (Figure 4) were assessed from the parameters of Model Ba and Model Bu (see Supplemental Table, Supplemental Digital Content 3, which gives the estimates, <http://links.lww.com/MSS/C804>). Additionally, computer simulations were done to determine the duration and rate of training reduction during taper which would give the highest performance according to the form of the taper, *i.e.* with step, linear or exponential training reduction, following previous reports (23-25). We computed optimal characteristics for a taper after regular training consisting in daily training equal to median TL ( $24.5 \pm 2.5$  training units) assumed to be long enough to stabilize performance and followed or not by a 20% step increase in training for 28 days. The characteristics of the optimal taper according to the form of training reduction are given in Table 2.

### **Validation datasets**

Figure 3 panel B shows prediction performance for Model Ba for one representative swimmer. The range of predicted values was narrower than that of real values with substantial underestimation and overestimation of the highest and the lowest performances respectively. Table 3 gives the systematic and typical errors for the group of swimmers. For the first validation data sets, typical error was greater than with the training data set, and without significant systematic error. For the second testing data sets, each model led to significant systematic error and a typical error greater than with training data sets. MAPE between predicted and real performance for the first data set was  $2.02 \pm 0.65$  % and  $2.17 \pm 0.65$  % for Model Ba and Model

Bu respectively. Greater MAPE was observed for the second testing data set,  $2.69 \pm 1.23$  % and  $2.56 \pm 0.79$  % for Model Ba and Model Bu respectively.

## DISCUSSION

The present findings are that Model Ba and Model Bu were relevant for modelling the change in performance with training in athletes, and showed a somewhat lesser ability to predict individual future performance from past data in a given swimmer. Impulse-response models suffer, however, from some limitations indicated in a series of recent studies (3, 9-11).

Model parameter estimation requires a great number of data points from valid, reliable and frequent tests of performance. We paid great attention to meeting these key requirements, leading us to use a test mimicking competitive racing conditions twice a week over an extended period. The results of the four trials done the first two week of this study showed a high reliability of this test. The coefficient of variation of  $0.40 \pm 0.23$  % was close to the value around 0.9% observed for field tests of sprint running (26). One limitation is the existence of local minima resulting in convergence to a solution sensitive to the initial parameters chosen for the starting point of the algorithm (10, 11). In our opinion, this troublesome problem is particularly relevant when all the parameters are estimated together using a typical fitting process. Finding initial estimates for the multiplying factors, i.e.  $kin_{pe}$ ,  $kin_{ne}$  or  $kin_{in_{ne}}$  depending on the model, is an awkward task because their estimates are dependent on the metrics used to quantify training and performance. This is the reason why we chose a hybrid approach to searching for parameter estimates, after checking that the solution identified was not sensitive to the starting point of the search for any of the models tested with our data. This hybrid approach

used a quasi-Newton algorithm, but only for a part of the parameters whereas the other parameters, *i.e.*  $kin_{pe}$ ,  $kin_{ne}$  or  $kin_{in_{ne}}$ , were estimated from simple regression at each step of the searching process.

Another important limitation is that it is possible to find different solutions which give close goodness-of-fit scores. This could preclude the use of model parameters as characteristics of one individual's response to training. We speculate that this might be due to compensation between model parameters which when combined together could give similar performance impulse responses. To support this assertion, we searched for other solutions than those used in the main study. We have repeated all the analyses with more restrictive constraints for  $kout_{pe}$  in order to find an alternative solution with much lower estimates for  $kout_{pe}$ . This resulted in large variations of the estimates of other parameters for each model (see document, Supplemental Digital Content 3, which gives details on these alternative solutions and their influence on model estimates, <http://links.lww.com/MSS/C804>). This alternative solution gave an optimized RSS close to that of the main studies. For Model Ba and Model Bu, the alternative sets of parameters gave small difference for the characteristics of the impulse response, although statistically significant differences were observed. As a consequence, results on optimal taper and prediction accuracy of performance are not significantly affected by this alternative solution. Because of the uncertainties of each model parameter, taken individually, it would be preferable to consider the impulse response obtained for an entire set of parameters when characterizing the response of one individual to training. In any case, this drawback regarding the estimation of model parameters did not appear to significantly affect the findings of the present study.

The impulse-response models tested in this study differ in (1) the description of positive and negative effects of a single training bout and (2) their consideration, or not, of change in impulse-response due to repetition of training bouts. One interest of indirect-response models lies in the differentiation of cumulated fatigue acting directly on performance and altering the athlete's capacity to adapt to the training. This can be particularly relevant when analyzing athlete behavior during overtraining (27). Indirect-response models have only been tested in participants in a controlled experiment who were not experienced athletes (13). In this his earlier study, Model TI outperformed other models. In the present data from swimmers, the TI model failed to effectively describe their variations in performance, and should be thus discarded for data analysis in athletes. When negative effect of training (*i.e.* fatigue) are considered, indirect response models fitted the data better than Model TI but did not better than Model Ba or Model Bu. AIC showed low weight of evidence in favor of Model TF, TF2 and TIF. Because of overfitting resulting in uncertainty in their estimates, these indirect-response models with fatigue components should be also discarded.

The absence of a clear difference in performance between Model Ba and Model Bu is in accordance with previous findings in athletes. On one hand, Model Bu was found to significantly improve performance fit compared to Model Ba in 7 out of 10 swimmers during a training period of 15 weeks (20). On the other hand, a study in 3 elite swimmers found higher  $R^2$  and lower AIC for Model Ba, compared to Model Bu (16). The authors pointed out that the frequency of performance tests could explain these discrepancies. They dispose of a limited number of performances in their data sets because they used results from competitions during a period of 400 days. The results of the present study also support this hypothesis when using only the first



20 weeks *i.e.* using only 40 values for performance. In this case, Adj.R<sup>2</sup> was significantly greater for Model Ba than Model Bu,  $0.36 \pm 0.09$  vs  $0.30 \pm 0.09$  respectively. In our previous report (20) and in the present study, we used performance tests done routinely during training sessions once and twice a week respectively. Results of competitions is the gold standard to determine athlete performance but the low number of competitions limits the statistical power when comparing modeling methods. This is why, despite possible bias, trials frequently repeated during the training sessions are an alternative which enables collection of large data sets on which to train and test models. Since the number of performance measurements was a major limit, it would be preferable to use Model Ba for data sets smaller than that of the present study, because it is the most parsimonious model.

Frequency of performance trials higher than in previous reports provided a larger amount of information which could be fruitful for exploring models more complex than Model Ba. In return, greater numbers of data points could affect fit quality. While R<sup>2</sup> values reported in the literature (28) ranged from 0.29 to 0.97, R<sup>2</sup> values for Model Ba and Model Bu obtained with the present data set were in the lower part of this range. The difference between model values and real performances showed that fit error was not randomly distributed. It is likely that the lower fit quality with a greater number of points was more due to greater loss of information than to greater noise. This suggests that Model Ba and Model Bu are not optimal models. Further work is necessary to find models that can better approximate the present large data sets, in the sense that models that loose less may serve to gain new insights into performance over time in adaptation to training.

Models could be used to probe the responses to solicitations and compare the impact of different training strategies (29, 30). Estimates of  $t_n$  and  $t_g$  have been considered to be key values for optimizing tapering period before a competition (31). According to Model Ba, taper duration should be ranged between  $t_n$  and  $t_g$  because negative effect (fatigue) dissipates more quickly than positive effect (adaptation or fitness) (31), *i.e.* between  $11.0 \pm 1.9$  and  $21.2 \pm 2.7$  days respectively in the present study. The major flaw in this reasoning is that the best performance would be indeed achieved without training during at least  $t_n$  days. It is recognized that several weeks without training would be prejudicial for performance because of detraining. Furthermore, the large estimates of  $t_n$ , which is in theory the time needed to recover performance after any training session, are not supported by the temporal variation of several physiological parameters altered by exercise (32). Contrary to Model Ba, Model Bu attempts to account for change in the response to a given workout among cumulated training sessions. A designated session could be more difficult to cope with when training is intensified. Conversely, the tolerance to training is recovered when training is reduced. This is illustrated by Fig. 4 which shows single bout and last bout effects at the asymptote calculated from the estimates of Model Bu. Mean recovery time for median TL would be ranged between 1 and 6 days depending on prior training. Compared to Model Ba, these estimates from Model Bu are in better agreement with within-subject variability in the time required to restore physiological alterations after working out (32). Additionally, computer simulations using estimates of Model Bu were done to explore the duration, the form and extent of training reduction to maximize performance during the taper period (23-25, 33, 34). Individual responses to training were simulated for accustomed training, evaluated as the median of TL, and for unaccustomed training as a 20% step increase in TL for 28 days. Contrary to Model Ba, estimates of Model Bu gave an optimal taper for around one week with a training

reduction of 80-90 % for a step taper and 60-70% for a progressive taper. This matches with an earlier review of the literature, which suggests a training reduction of 50 – 90% over 4 – 35 days regardless the form of the training reduction (35). It agrees less with a meta-analysis study which showed that the most efficient taper for maximizing performance in swimmers was 1 or 2 weeks with a progressive reduction of training volume by 40-60%, without any modification of intensity (36). Contrary to our previous report in elite swimmers (23), greater solicitations before the taper require only a small increase in taper duration which remains around 1 week and progressive reduction of training did not give better results than step reduction. These differences could be explained by the lower level of performance of the participants in the present study. Whereas our previous report in best nationally ranked swimmers (23) concluded that optimal taper would be a training reduction of 40 to 70 % during 2 to 3 weeks, depending on prior training, the present findings suggest for sub-elite swimmers shorter taper lasting around 1 week with greater training reduction, around 60 to 90 %.

In our analysis, we examined how well models predicted the future from past data and compared model performance over different time spans. The theoretical studies on taper characteristics based on model predictions (23-25, 33, 34) has to be distinguished from using one model to monitor training of a given individual because this kind of application would require a more precise assessment of the statistical confidence of the model prediction (37). Prediction accuracy was lower than in our previous report in swimmers (20). It could be explained by the closeness of the training cycles used to train the models and to test their prediction. It is the reason why we shared the data to generate two validation data sets per swimmer: first one for testing the prediction during training cycle immediately following the training period used to

parametrize the models and a second one for testing the prediction for the subsequent season after the summer off-season. This design showed that the agreement in performance prediction dropped for the last cycle of Season 1 to the two cycles of Season 2 due to a large bias for Season 2. Typical errors indicated Model Ba performed better for Season 1, while Model Bu was better for Season 2 but the differences in the performance of the two models were too small to be considered to be significant. Altogether relative accuracy for predicting performance was around 2-3 % for both models, whereas the fluctuations of actual performance throughout the two seasons were between 3% below and 7% above the first value of Season 1 (Figure 2). This accuracy level does not effectively meet the need of athletes if models were to be implemented in practical application to assist coaches when planning training of a given individual. More accurate prediction is required before considering these models for such applications.

The literature provides other model structure solutions which should also be tested for their ability to fit the observed performances and ultimately to predict future performance. One important feature not included in Model Ba nor in Model Bu is that positive effects are linearly dependent on TL. This is a critical issue when attempting to better represent maladaptation because of overtraining. It is the reason why the analysis of optimal taper focused only on how reduce training for peak performance. The computations were done for regular training and for a step 20% increase for 28 days before taper in order to remain within the range of the training effort actually produced by swimmers. Optimal training before taper cannot be searched because the gain in performance after optimal taper is linearly related to the training done before the reduction of training, despite limitations in training adaptations. In the models tested in the present study, only Model TI and Model TIF included inhibitory processes in an attempt to

introduce limitations in the positive effect of training, and neither of them provided any improvement compared to Model Ba or Model Bu. Other modifications of Model Ba (38, 39) have been proposed to take into consideration that positive effect is not simply proportional to TL. It was showed that modification of Model Ba with Hill saturation function resulted in better fit in elite swimmers analyzed over a period of  $4 \pm 2$  years (38). In line with the indirect-response models tested in this study, there are other formulations using delayed positive effect after training session (8, 14, 15, 40). Using two exponentials for positive effect was found to outperform Model Bu with data sets from short-track sprint skaters (15). Extensions of Model Ba or Model Bu have been also formulated to account for prior training history (41) or temporal variation in model parameters (42, 43). Methodologies other than impulse-response models also deserve attention: Per-Pot model (44), neural networks (45, 46), exponentially weighed and rolling average models (16) and machine-learning models (4, 5). One way to assess the relative strengths and limitations of these different models is to compare their performance using a common data set. This is why we believe that the database generated for the present study has considerable potential for reuse. It is rich enough to support testing of a wide range of models that predict performance from cumulated training. Nevertheless, the strongest limitation of these models is probably the use of one single metric for quantifying TL as model input (37). It does not allow us to take into consideration the specificities of the different types of exercises included in the training program. For example, variations in volume or intensity could give identical TL despite difference in their effects. Regardless of the variation in TL for peaking performance, how the content of sessions is determined, and how they are articulated over time, is a great part of the coach's work when planning training. More complex models or combination of different approaches are most likely needed to account for all the key factors in training

planning.

## **CONCLUSIONS**

In conclusion, this study showed that Model Ba and Model Bu were equally relevant for modelling the change in performance with training in athletes. Because of the lower number of adjustable parameters, Model Ba should be preferred to Model Bu to provide information about response to training. The interest in using Model Bu could arise from the inferences made from its application to dedicated large data sets, rather than precise prediction for a given individual. The observed accuracy of 2 to 3% for the prediction of one individual's performance in response to training was not satisfactory. The challenge to progress towards practical applications to assist coaches and athletes for planning training requires exploration of other approaches to modeling the responses to training.

## **Acknowledgments**

The authors gratefully thank all the swimmers who participated in this study. The authors also acknowledge Crane Rogers for editing the manuscript for English language and helpful discussion.

## **Conflict of Interest**

The authors declare that they have no conflicting interests. The results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. The results of the present study do not constitute endorsement by the American College of Sports Medicine.

ACCEPTED

## REFERENCES

1. Rasche C, Pfeiffer M. Training. In: A Baca, J Perl editors. *Modelling and Simulation in Sport and Exercise*. Routledge; 2018, pp. 187-207.
2. Stephens Hemingway B, Greig L, Jovanovic M, Ogorek B, Swinton P. Traditional and contemporary approaches to mathematical fitness-fatigue models in exercise science: A practical guide with resources. Part I. [Internet]. SportRxiv; 2021. Available from: [osf.io/ap75j](https://osf.io/ap75j) 2020.
3. Swinton P, Stephens Hemingway B, Rasche C, Pfeiffer M, Ogorek B. Traditional and contemporary approaches to mathematical fitness-fatigue models in exercise science: a practical guide with resources. Part II. [Internet]. SportRxiv; 2021. Available from: [osf.io/5qgc2](https://osf.io/5qgc2)
4. Imbach F, Perrey S, Chailan R, Meline T, Candau R. Training load responses modelling and model generalisation in elite sports. *Sci Rep*. 2022;12(1):1586.
5. Imbach F, Sutton-Charani N, Montmain J, Candau R, Perrey S. The use of fitness-fatigue models for sport performance modelling: conceptual issues and contributions from machine-learning. *Sports Med Open*. 2022;8(1):29.
6. Mujika I, Padilla S, Pyne D, Busso T. Physiological changes associated with the pre-event taper in athletes. *Sports Med*. 2004;34(13):891-927.
7. Banister EW, Calvert TW, Savage MV, Bach T. A systems model of training for athletic performance. *Aust J Sports Med*. 1975;7:57-61.
8. Calvert TW, Banister EW, Savage MV, Bach T. A systems model of the effects of training on physical performance. *IEEE Trans Syst Man Cybern*. 1976;6:94-102.
9. Stephens Hemingway BH, Burgess KE, Elyan E, Swinton PA. The effects of measurement error and testing frequency on the fitness-fatigue model applied to resistance



training: a simulation approach. *Int J Sports Sci Coach*. 2020;15(1):60-71.

10. Stephens Hemingway B, Swinton P, Ogorek B. The suitability of a quasi-Newton algorithm for estimating fitness-fatigue models: Sensitivity, troublesome local optima, and implications for future research (an in silico experimental design). [Internet]. SportRxiv; 2021. Available from: [osf.io/dx7gm](https://osf.io/dx7gm).
11. Vermeire K, Ghijs M, Bourgeois JG, Boone J. the fitness–fatigue model: what’s in the numbers? *Int J Sports Physiol Perform*. 2022;17(5):810-3.
12. Busso T. Variable dose-response relationship between exercise training and performance. *Med Sci Sports Exerc*. 2003;35(7):1188-95.
13. Busso T. From an indirect response pharmacodynamic model towards a secondary signal model of dose-response relationship between exercise training and physical performance. *Sci Rep*. 2017;7:40422.
14. Philippe AG, Borrani F, Sanchez AM, Py G, Candau R. Modelling performance and skeletal muscle adaptations with exponential growth functions during resistance training. *J Sports Sci*. 2019;37(3):254-61.
15. Borrani F, Solsona R, Candau R, Meline T, Sanchez AM. Modelling performance with exponential functions in elite short-track speed skaters. *J Sports Sci*. 2021;39(20):2378-85.
16. Mitchell LJG, Rattray B, Fowlie J, Saunders PU, Pyne DB. The impact of different training load quantification and modelling methodologies on performance predictions in elite swimmers. *Eur J Sport Sci*. 2020;20(10):1329-38.
17. Burnham KP, Anderson DR. *Model Selection and Multimodel Inference*. 2nd ed. New York: Springer; 2002, 488 p.
18. Busso T, Carasso C, Lacour JR. Adequacy of a systems structure in the modeling of

training effects on performance. *J Appl Physiol* (1985). 1991;71(5):2044-9.

19. Spiess AN, Neumeyer N. An evaluation of R<sup>2</sup> as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach. *BMC Pharmacol*. 2010;10:6.
20. Chalencon S, Pichot V, Roche F, et al. Modeling of performance and ANS activity for predicting future responses to training. *Eur J Appl Physiol*. 2015;115(3):589-96.
21. Mujika I, Busso T, Lacoste L, Barale F, Geysant A, Chatard JC. Modeled responses to training and taper in competitive swimmers. *Med Sci Sports Exerc*. 1996;28(2):251-8.
22. R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
23. Thomas L, Mujika I, Busso T. A model study of optimal training reduction during pre-event taper in elite swimmers. *J Sports Sci*. 2008;26(6):643-52.
24. Sanchez AM, Galbes O, Fabre-Guery F, et al. Modelling training response in elite female gymnasts and optimal strategies of overload training and taper. *J Sports Sci*. 2013;31(14):1510-9.
25. Meline T, Mathieu L, Borrani F, Candau R, Sanchez AM. Systems model and individual simulations of training strategies in elite short-track speed skaters. *J Sports Sci*. 2019;37(3):347-55.
26. Hopkins WG, Schabert EJ, Hawley JA. Reliability of power in physical performance tests. *Sports Med*. 2001;31(3):211-34.
27. Meusen R, Duclos M, Foster C, et al. Prevention, diagnosis, and treatment of the overtraining syndrome: joint consensus statement of the European College of Sport Science and the American College of Sports Medicine. *Med Sci Sports Exerc*. 2013;45(1):186-205.
28. Pfeiffer M. Modeling the relationship between training and performance - a comparison

of two antagonistic concepts. *Int J Comput Sci Sport*. 2008;7(2):13-32.

29. Clarke DC, Skiba PF. Rationale and resources for teaching the mathematical modeling of athletic training and performance. *Adv Physiol Educ*. 2013;37(2):134-52.

30. Morton RH. The quantitative periodization of athletic training: a model study. *Sports Med Training Rehabilitation*. 1991;3:19-28.

31. Fitz-Clarke JR, Morton RH, Banister EW. Optimizing athletic performance by influence curves. *J Appl Physiol (1985)*. 1991;71(3):1151-8.

32. Skorski S, Mujika I, Bosquet L, Meeusen R, Coutts AJ, Meyer T. The temporal relationship between exercise, recovery processes, and changes in performance. *Int J Sports Physiol Perform*. 2019;14(8):1015-21.

33. Thomas L, Busso T. A theoretical study of taper characteristics to optimize performance. *Med Sci Sports Exerc*. 2005;37(9):1615-21.

34. Thomas L, Mujika I, Busso T. Computer simulations assessing the potential performance benefit of a final increase in training during pre-event taper. *J Strength Cond Res*. 2009;23(6):1729-36.

35. Mujika I, Padilla S. Scientific bases for precompetition tapering strategies. *Med Sci Sports Exerc*. 2003;35(7):1182-7.

36. Bosquet L, Montpetit J, Arvisais D, Mujika I. Effects of tapering on performance: a meta-analysis. *Med Sci Sports Exerc*. 2007;39(8):1358-65.

37. Busso T, Thomas L. Using mathematical modeling in training planning. *Int J Sports Physiol Perform*. 2006;1(4):400-5.

38. Hellard P, Avalos M, Millet G, Lacoste L, Barale F, Chatard JC. Modeling the residual effects and threshold saturation of training: a case study of Olympic swimmers. *J Strength Cond*

*Res.* 2005;19(1):67-75.

39. Turner JD, Mazzoleni MJ, Little JA, Sequeira D, Mann BP. A nonlinear model for the characterization and optimization of athletic training and performance. *Biomed Hum Kinet.* 2017;9:82-93.
40. Matabuena M, Rodriguez-Lopez R. An improved version of the classical Banister model to predict changes in physical condition. *Bull Math Biol.* 2019;81(6):1867-84.
41. Ludwig M, Asteroth A, Rasche C, Pfeiffer M. Including the past: performance modeling using a preload concept by means of the fitness-fatigue model. *Int J Comput Sci Sport.* 2019;18(1):115-34.
42. Busso T, Denis C, Bonnefoy R, Geysant A, Lacour JR. Modeling of adaptations to physical training by using a recursive least squares algorithm. *J Appl Physiol (1985).* 1997;82(5):1685-93.
43. Kolossa D, Bin Azhar MA, Rasche C et al. Performance estimation using the fitness-fatigue model with Kalman filter feedback. *Int J Comput Sci Sport.* 2016;16(2):117-29.
44. Perl J. PerPot: a metamodel for simulation of load performance interaction. *Eur J Sport Sci.* 2001;1(2):1-13.
45. Edelmann-nusser J, Hohmann A, Henneberg B. Modeling and prediction of competitive performance in swimming upon neural networks. *Eur J Sport Sci.* 2002;2(2):1-10.
46. Carrard J, Kloucek P, Gojanovic B. Modelling training adaptation in swimming using artificial neural network geometric optimisation. *Sports (Basel).* 2020;8(1):8.

## FIGURE LEGENDS

**Figure 1:** Schematic representation of impulse responses of models. A: positive and negative effects of a single session, both being described by a single exponential with Model Ba. B : performance over time based on the difference between positive and negative effects with Model Ba,  $\Delta P_0$  : decrease in performance after the training session,  $t_n$  : time needed to recover to the initial performance level  $p_g$  : maximal gain in performance and  $t_g$  : time needed to reach  $p_g$ . C : change in negative effect between response to an isolated training bout (single session) and to last bout after several repetitions (Model Bu). D : change in performance over time between single session and repetition (Model Bu). E : positive effect described by a single exponential ( Models Ba and Bu) vs secondary signal model ( Model TI, TF, TF2 and TIF). F : positive effect by secondary signal model with inhibition due to previous sessions ( Model TI and TIF).

**Figure 2:** Distribution of training loads and change in performance. A and B : Mean  $\pm$  SD of weekly training loads (TL) during season 1 and 2 respectively. C and D : Mean  $\pm$  SD of performance relative to first value during season 1 and 2 respectively.

**Figure 3:** Modeling and prediction of performance for one representative swimmer with Model Ba. A : model fit to all the data of Season 1 and Season 2. B : model fit to the data of cycles # 1 and 2 (training data set) and prediction tested using data of cycles # 3, 4 and 5 (testing data sets). Solid circle : observed performance. Continuous line : modelled performance.

**Figure 4:** Mean  $\pm$  SD of characteristics of impulse response to median TL from estimates of Model Ba and Bu. A : change in performance after one session. B :  $\square P_0$ , initial decrease in performance. C :  $p_g$ , maximal gain in performance. D :  $t_n$ , time needed to recover to initial performance level. E :  $t_g$  : time needed to reach  $p_g$ . For model Bu, results show the response to an isolated training bout (single session) and to the last bout at steady state after daily repetitions (Asymptote). \* : significantly different from Model Ba and †: significantly different from Model Bu for single session ( $P < 0.05$ ).

#### **SUPPLEMENTAL DIGITAL CONTENT**

**SDC 1:** Supplemental Digital Content 1. pdf

**SDC 2:** Supplemental Digital Content 2. pdf

**SDC 3:** Supplemental Digital Content 3. pdf

Figure 1

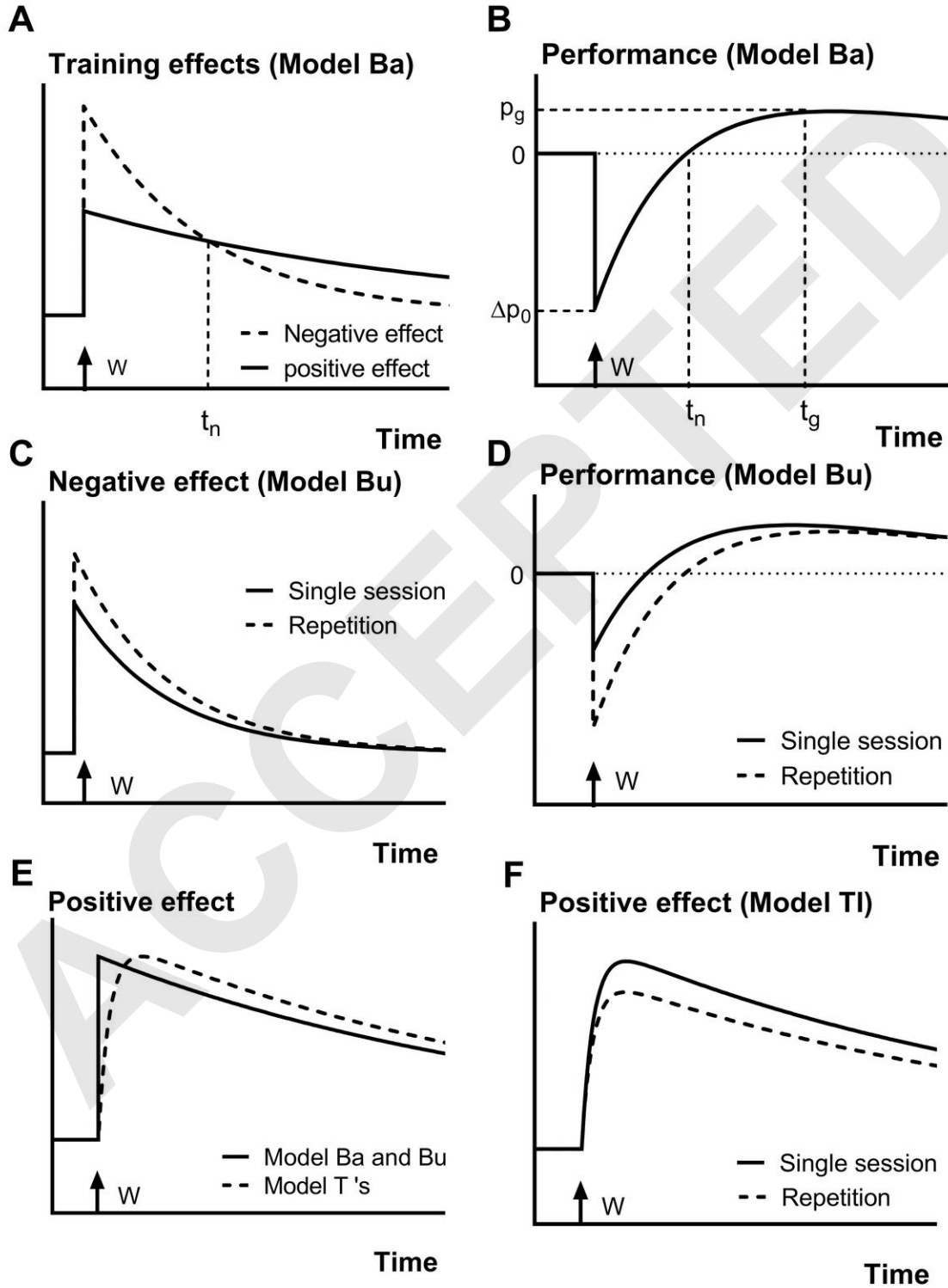


Figure 2

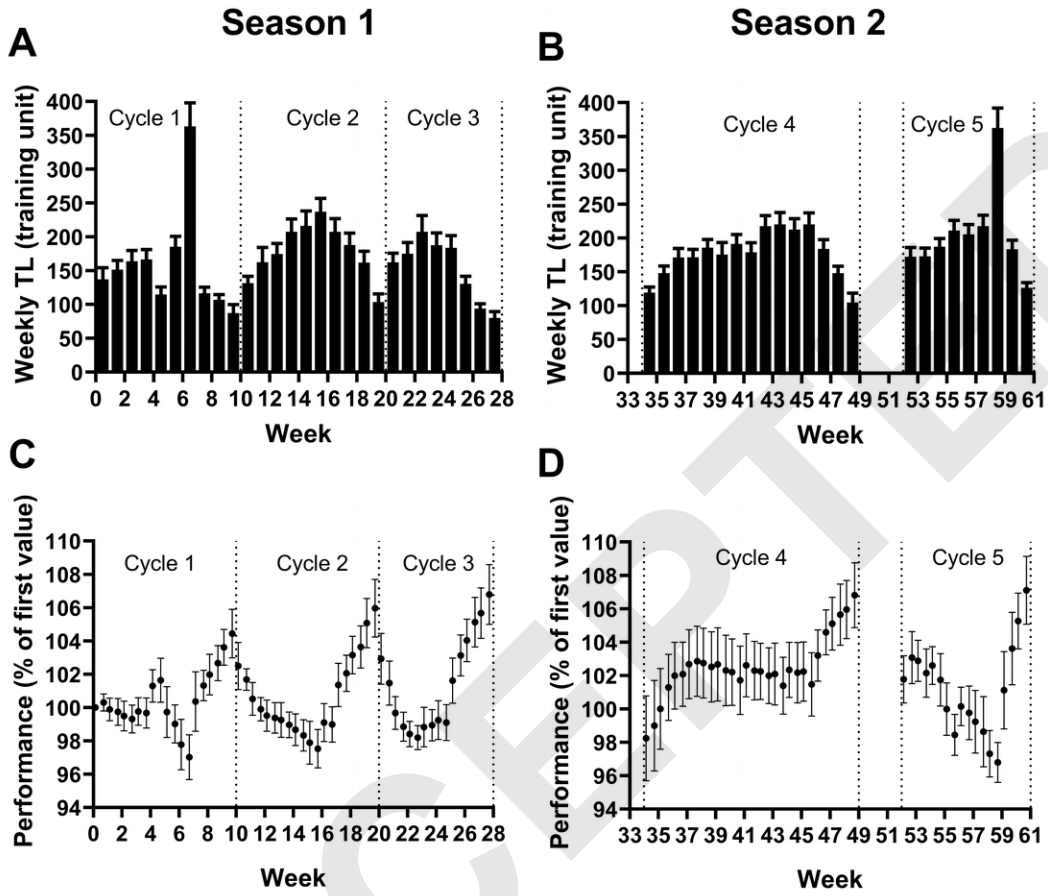




Figure 3

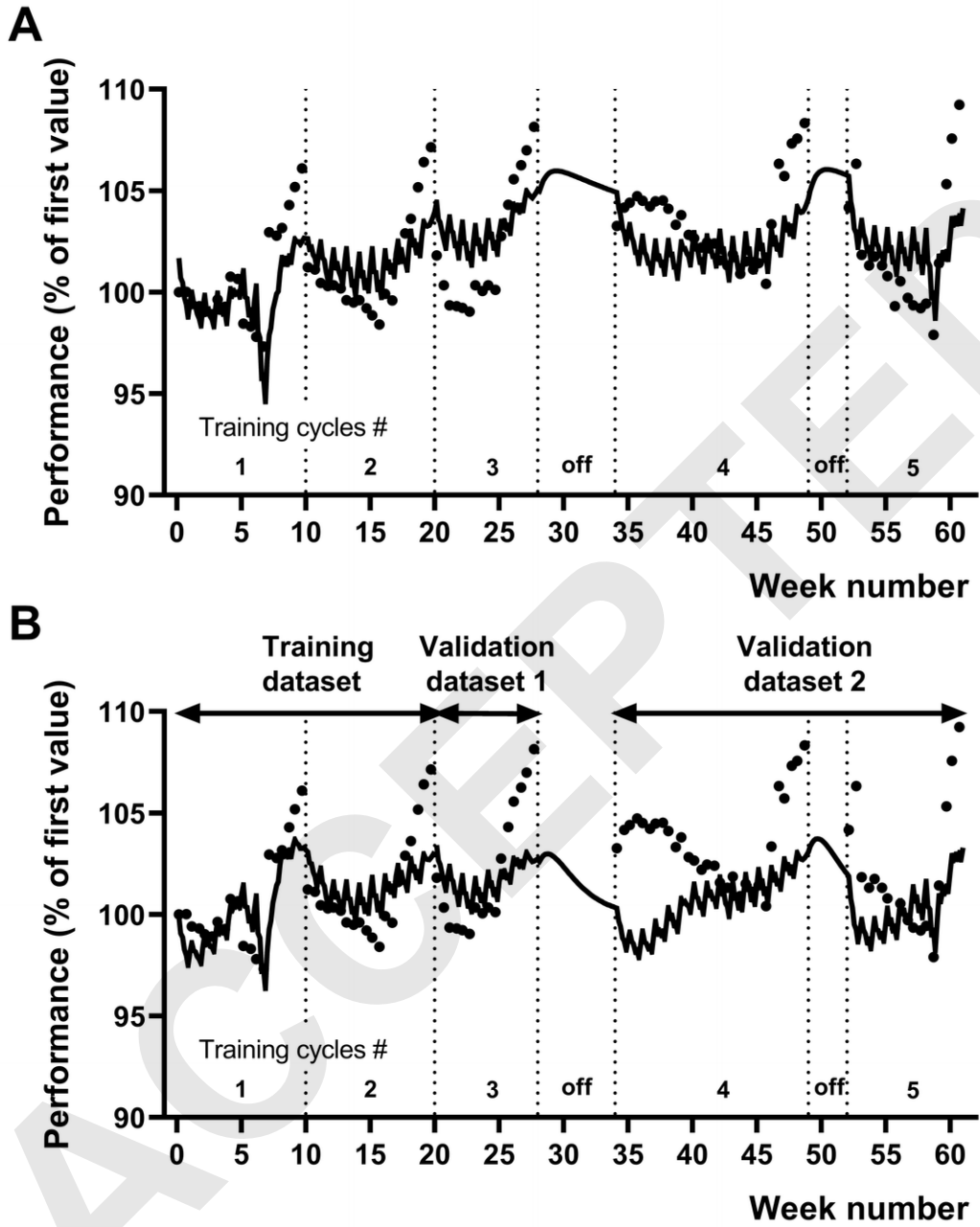
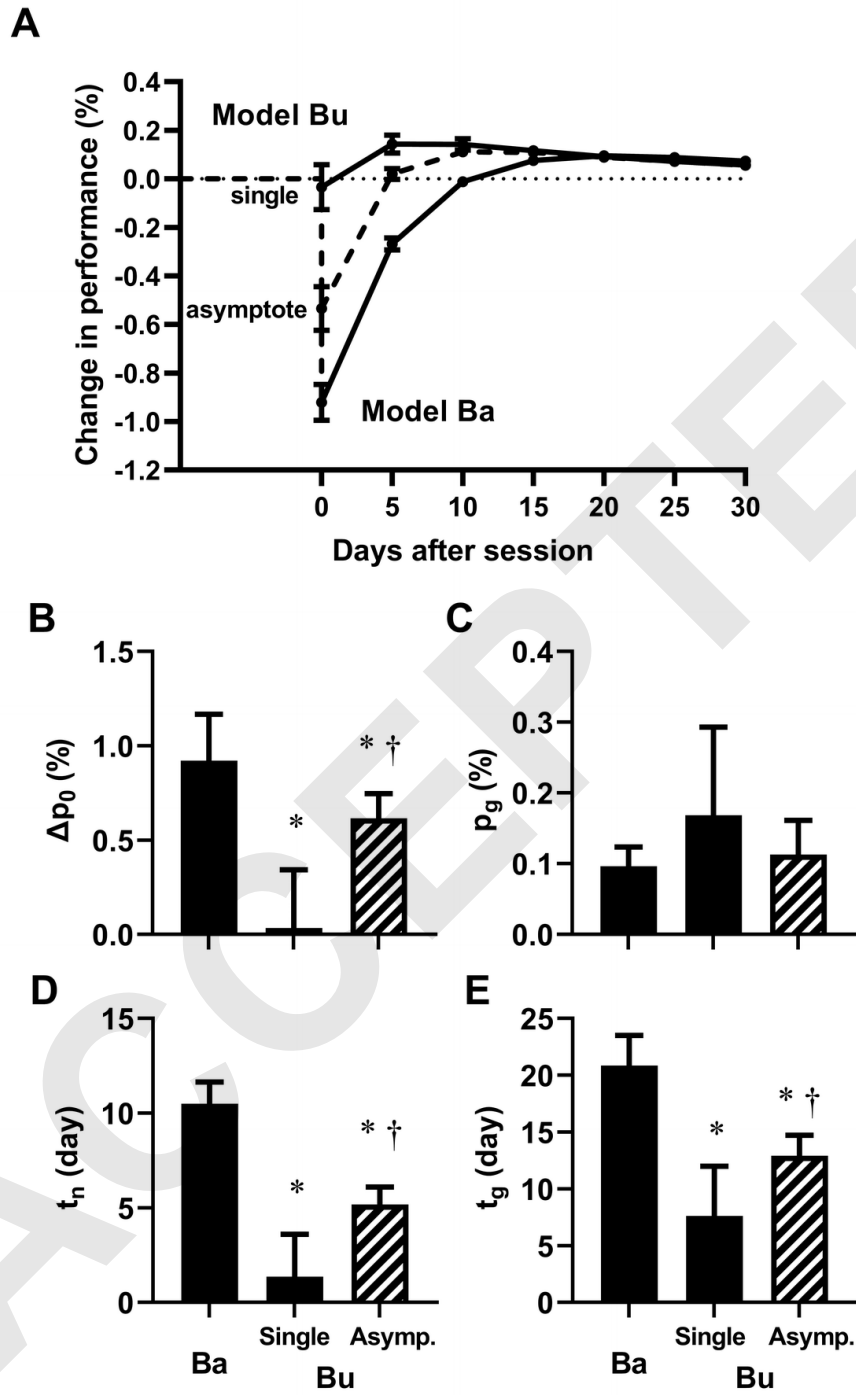


Figure 4



**Table 1.** Summary of goodness-of-fit and AIC results for models with all data of Seasons 1 and 2.

Model	R <sup>2</sup>	adj-R <sup>2</sup>	SE	MAPE (%)	AICc	w(AIC)
Ba (p=5)	0.286 ± 0.100	0.250 ± 0.097	2.28 ± 0.53	1.82 ± 0.45	409 ± 112	0.339 ± 0.254
Bu (p=6)	0.297 ± 0.073	0.252 ± 0.071	2.28 ± 0.52	1.80 ± 0.42	409 ± 109	0.360 ± 0.296
TI (p=5)	0.135 * † ± 0.053	0.091 * † ± 0.062	2.52 * † ± 0.56	2.00 * † ± 0.44	425 * † ± 108	0.001 * † ± 0.002
TF (p=6)	0.288 ‡ ± 0.098	0.243 * ‡ ± 0.094	2.29 * ‡ ± 0.53	1.81 ‡ ± 0.45	411 * ‡ ± 112	0.115 * ‡ ± 0.082
TF2 (p=7)	0.304 ‡ ± 0.081	0.251 ‡ ± 0.077	2.29 ‡ ± 0.53	1.79 ‡ ± 0.43	411 † ‡ ± 110	0.123 ‡ ± 0.088
TIF (p=7)	0.299 * ‡ ± 0.097	0.246 ‡ ± 0.091	2.29 ‡ ± 0.53	1.80 * ‡ # ± 0.45	412 ‡ ± 112	0.061 * ‡ ± 0.036

Values are mean ± SD. p: number of model parameters. Number of performance points is 56 for 3 swimmers (Season 1) and 104 for 8 swimmers (Season 1 and 2). Significant difference (P<0.05) indicated by \*: from Model Ba, †: from Model Bu, ‡: from Model TI, #: from Model TF.

**Table 2 :** Mean  $\pm$  SD of estimated characteristics of the optimal simulated taper according to the form of the training reduction after regular training at the median TL and overload training with a step increase of 20% for 28 days.

	Form of the training reduction during taper		
	Step	Linear	Exponential
<b>Mean reduction (%)</b>			
Regular training	86.9 $\pm$ 16.0	73.5 $\pm$ 12.3 †	61.5 $\pm$ 12.6 †
Regular training + 20%	86.3 $\pm$ 13.8	73.3 $\pm$ 8.8 †	66.4 $\pm$ 12.6 †
<b>Duration (days)</b>			
Regular training	6.5 $\pm$ 1.2	9.0 $\pm$ 1.1 †	8.2 $\pm$ 0.6 †
Regular training + 20%	7.6 $\pm$ 0.7 *	9.5 $\pm$ 1.7 * †	8.7 $\pm$ 1.0 * †
<b>Highest performance (% of pre-training)</b>			
Regular training	103.24 $\pm$ 1.12	103.21 $\pm$ 1.13 †	103.21 $\pm$ 1.13 †
Regular training + 20%	103.85 $\pm$ 1.30 *	103.82 $\pm$ 1.31 * †	103.83 $\pm$ 1.30 * †

\* : statistically different from regular training prior taper ( P<0.05).

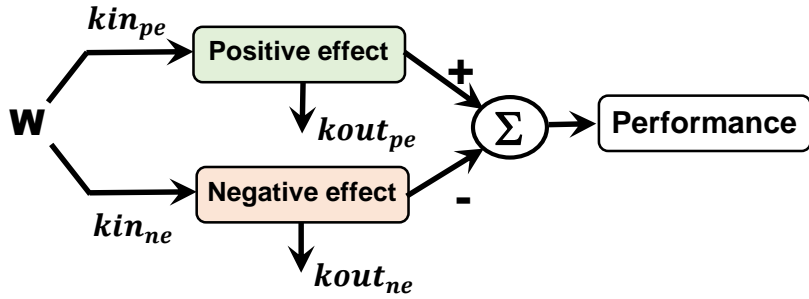
† : statistically different from step taper (P<0.05).

**Table 3.** Reliability of the performance predictions.

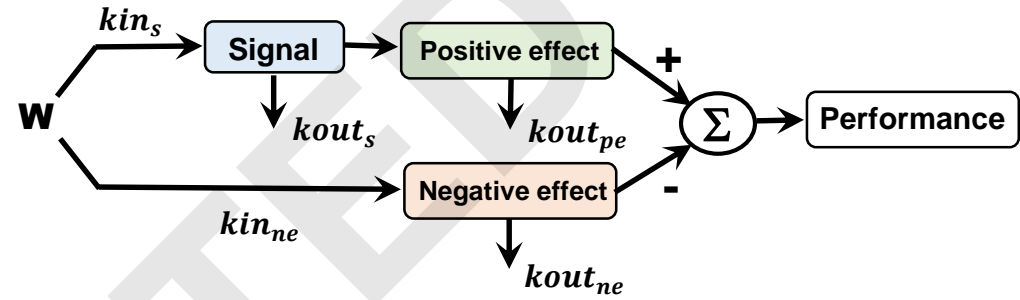
Model	Data set	Systematic Error (%)	Typical Error (%)	MAPE (%)
Ba	Training	0	1.68 ± 0.45	1.34 ± 0.37
	Validation 1	0.16 ± 0.40	2.40 ± 0.74 <sup>a</sup>	2.02 ± 0.65 <sup>a</sup>
	Validation 2	-1.74 ± 1.23 <sup>a,b</sup>	2.43 ± 0.51 <sup>a</sup>	2.69 ± 1.23 <sup>a,b</sup>
Bu	Training	0	1.73 ± 0.47	1.39 ± 0.39
	Validation 1	0.21 ± 0.42	2.56 ± 0.80 <sup>a,*</sup>	2.17 ± 0.70 <sup>a</sup>
	Validation 2	-1.64 ± 1.15 <sup>a,b</sup>	2.32 ± 0.49 <sup>a,*</sup>	2.56 ± 0.79 <sup>a,b</sup>

Values are mean ± SD. Systematic and typical error are expressed as a percentage of first performance of Season 1. Significant difference (P<0.05) indicated by <sup>a</sup> : from Training data set, <sup>b</sup> : from first validation data set, \*: from Model Ba.

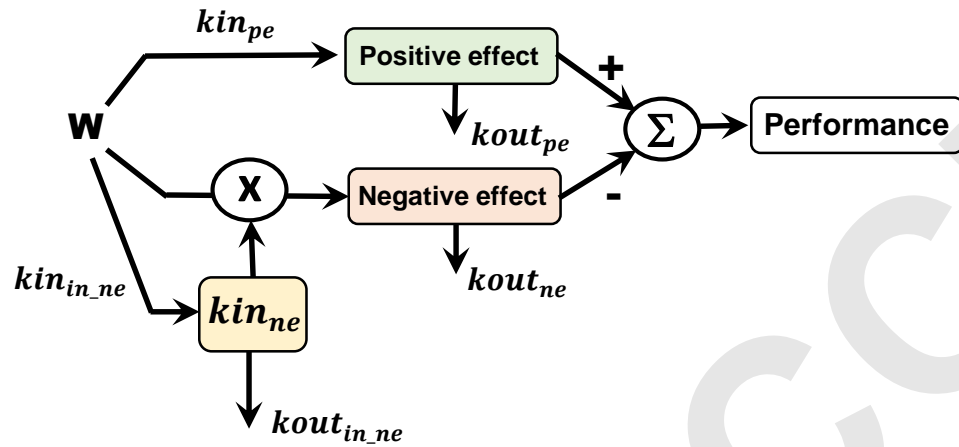
Model Ba



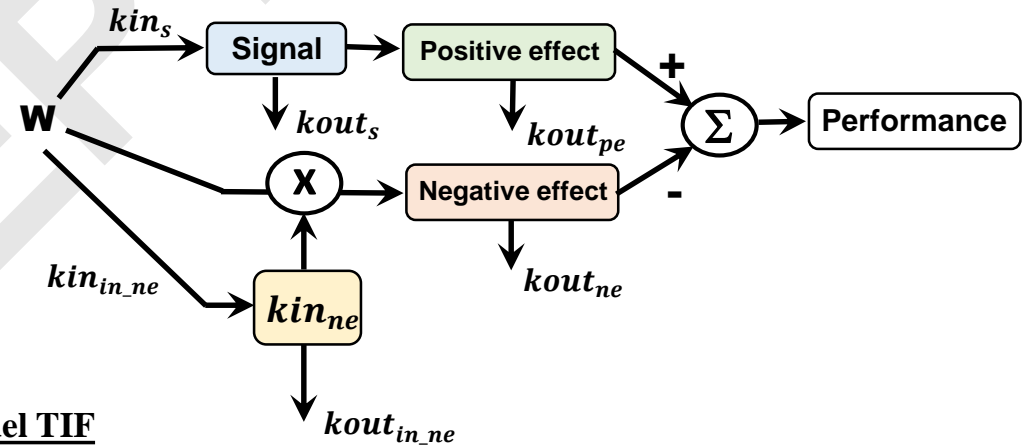
Model TF



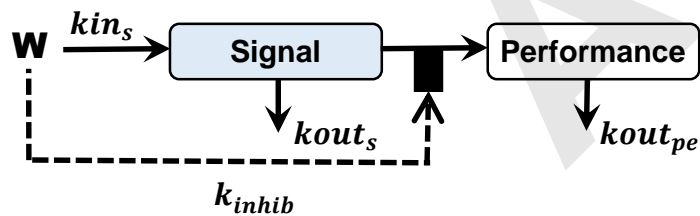
Model Bu



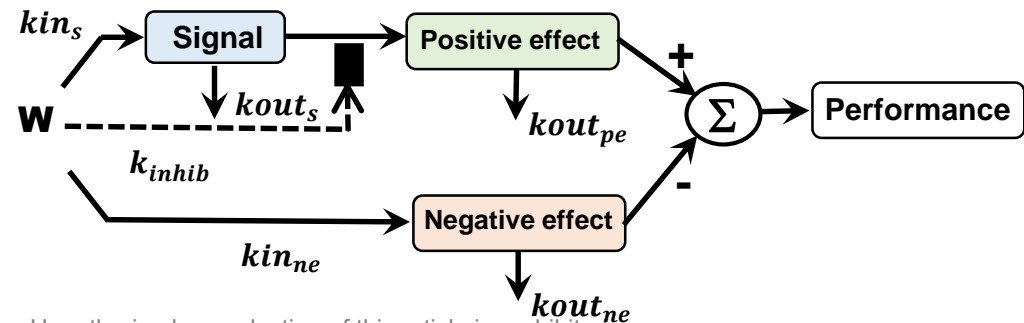
Model TF2



Model TI



Model TIF



```

Fit_Model<-function(datafile,model=c("Ba","Bu","TI","TF","TF2","TIF"))

{

# General purpose
#
# The function fits impulse-response models to data by minimizing residual squared errors using constrOptim function.
#
#
# Details
#
# The method for optimization (L-BFGS-B) is that of Byrd et al (1995) which is a modification of
# BFGS quasi-Newton method which allows box constraints.
# Lower and upper bounds and initial values satisfying the constraints are given for each parameter.
#
#
# Arguments
#
# "datafile" is the file containing the data
# column 1 : day number incremented by one starting from one
# column 2 : training load of the day, zero if no training
# column 3 : measure of performance of the day, empty if no measurement
#
# "model" is the model which is fitted to the data
#
# Ba : model with two exponentials (Banister et al, 1975)
# Bu : model with mutiplying factor for fatigue dependent on training loads (Busso 2003)
# TI : model with indirect secondary signal and inhibition of positive effect (Busso 2017)
# TF : model with indirect secondary signal and fatigue as in Ba (Busso 2017)
# TF2 : model with indirect secondary signal and fatigue as in Bu (unpublished)
# TIF : model with indirect secondary signal, inhibition and fatigue as in Ba (Busso 2017)
#
#
# Values
#

```

```

# Measures of goodness-of-fit :
# r2, coefficient of determination
# r2adj, adjusted coefficient of determination
# logL, maximum log-likelihood
# aic, Akaike Information Criterion
# aicc, bias-corrected Akaike Information Criterion
# se, mean square error
#
# "converg" is an integer code : "0" indicates successful completion of optimization
# (see description of Optim function for more details).
#
# Estimates of model paramaters :
# pbase : baseline value of performance
# kout_pe : exponential decay rate of positive effect (all models)
# kin_pe : multiplying factor of positive effect (Model Ba and Bu)
# kout_ne : exponential decay rate of negative effect (Model Ba, Bu, TF, TF2 and TIF)
# kin_ne : multiplying factor of negative effect (Model Ba, TF and TIF)
# kout_in_ne : exponential decay rate of kin_ne (Model Bu and TF2)
# kin_in_ne : multiplying factor of kin_ne (Model Bu and TF2)
# kout_s : exponential decay rate of secondary signal (Model TI, TF, TF2 and TIF)
# kin_s : multiplying factor secondary signal (Model TI, TF, TF2 and TIF)
# kinhib : multiplying factor for inhibition of positive effect (Model TI and TIF)

# Fixing model and its number of parameters, p

model<-match.arg(model)
if (model == "Ba") p<-5
if (model == "Bu") p<-6
if (model == "TI") p<-5
if (model == "TF") p<-6
if (model == "TF2") p<-7
if (model == "TIF") p<-7

# Fixing starting values and restrictions of model parameters for optimization process

```



```

kout_pe<-NA;kout_ne<-NA;kout_s<-NA;kout_in_ne<-NA;kinhib<-NA
kin_pe<-NA;kin_ne<-NA;kin_in_ne<-NA;kin_s<-NA

kout_pe_init<-0.03;kout_s_init<-0.1;kinhib_init<-0.001;kout_ne_init<-0.1;kout_in_ne_init<-0.5
kout_pe_min<-0.01;kout_s_min<-0.05;kinhib_min<-0;kout_ne_min<-0.06;kout_in_ne_min<-0.1
kout_pe_max<-0.07;kout_s_max<-1;kinhib_max<-0.01;kout_ne_max<-1;kout_in_ne_max<-2

# Fixing data from datafile given in the argument
# extract : days with performance measurement
# pmes : measured performance
# w : training load
# n : number of performance measurements
# ntot : total day number

extract<-datafile[1][is.finite(datafile[,3])]
w<-datafile[,2]
pmes<-datafile[,3]
n<-length(extract)
ntot<-length(w)

# Computations of sumw and sumw2 used in the functions called for compute training effects

toep_w<-toeplitz(c(0,w[1:ntot-1]))
sumw<-lower.tri(toep_w,diag=TRUE)*toep_w
toep_w2<-toeplitz(c(0,w[1:ntot-1]^2))
sumw2<-lower.tri(toep_w2,diag=TRUE)*toep_w2

# Functions called for computing training effects
#
# pe_1exp : computation of positive effect with Model Ba and Bu
# pe_T : computation of positive effect with Model TF and TF2
# pe_TI : computation of positive effect with Model TI and TIF
# ne_1exp : computation of negative effect with Model Ba, TF and TIF
# ne_2exp : computation of negative effect with Model Bu and TF2

```

```
pe_1exp<-function(kout_pe) sumw%%c(exp(-(1:ntot)*kout_pe))
```

```
ne_1exp<-function(kout_ne) sumw%%c(exp(-(1:ntot)*kout_ne))
```

```
ne_2exp<-function(kout_ne,kout_k2) {  
  kin2<-sumw%%c(exp(-(0:(ntot-1))*kout_k2))  
  toep_wk2<-toeplitz(c(0,w[1:(ntot-1)])*c(kin2))  
  sumwk2<-lower.tri(toep_wk2,diag=TRUE)*toep_wk2  
  ne_2exp<-sumwk2%%c(exp(-(1:ntot)*kout_ne))  
}
```

```
pe_T<-function(kout_pe,kout_s) {  
  sign<-sumw%%c(exp(-(0:(ntot-1))*kout_s))  
  toep_e<-toeplitz(c(sign))  
  sum_prod<-lower.tri(toep_e,diag=TRUE)*toep_e  
  pe_T<-sum_prod%%exp(-(0:(ntot-1))*kout_pe)  
}
```

```
pe_TI<-function(kout_pe,kout_s,kin_i) {  
  TI<-1-kin_i*w  
  sign<-sumw%%c(exp(-(0:(ntot-1))*kout_s))  
  prod<-sign*TI  
  toep_e<-toeplitz(c(prod))  
  sum_prod<-lower.tri(toep_e,diag=TRUE)*toep_e  
  pe_TI<-sum_prod%%exp(-(0:(ntot-1))*kout_pe)  
}
```

```
# Computations for Model Ba
```

```
if (model == "Ba"){
```

```
# Fixing starting values and restrictions of model parameters for optimization
```

```
init<-c(kout_pe_init,kout_ne_init)
```

```

mconstr<-rbind(diag(1,2,2),c(-1,0),c(0,-1))
vconstr<-c(c(kout_pe_min,kout_ne_min),c(-kout_pe_max),c(-kout_ne_max))

# Parameters of Model Ba are optimized by a two-stage minimization of the sum of squared errors,rss
# 1. pbase, kin_pe and kin_ne are computed by linear regression according to values for kout_pe and kout_ne
# 2. kout_pe and kout_ne are optimized with constrOptim function

rss<-function(param){
  kout_pe<-param[1]
  kout_ne<-param[2]
  datareg<-list(pe=pe_1exp(kout_pe)[extract],ne=ne_1exp(kout_ne)[extract],p=pmes[extract])
  reg<-lm(p~pe+ne,datareg)
  pmod<-reg$coefficients[1]+reg$coefficients[2]*pe_1exp(kout_pe)+reg$coefficients[3]*ne_1exp(kout_ne)
  neglogL<-0.5*n*(log(2*pi)+1-log(n)+log(sum((pmes-pmod[extract])^2)))
  rss<-sum((pmes[extract]-pmod[extract])^2)
}

fit<-constrOptim(theta=init,f=rss,grad=NULL,ui=mconstr,ci=vconstr,control=list(maxit=3000))
kout_pe<-fit$par[1]
kout_ne<-fit$par[2]
datareg<-list(pe=pe_1exp(kout_pe)[extract],ne=ne_1exp(kout_ne)[extract],p=pmes[extract])
reg<-lm(p~pe+ne,datareg)
pbase<-reg$coefficients[1]
kin_pe<-reg$coefficients[2]
kin_ne<-reg$coefficients[3]
pmod<-pbase+kin_pe*pe_1exp(kout_pe)-kin_ne*ne_1exp(kout_ne)
}

if (model == "Bu"){

# Fixing starting values and restrictions of model parameters for optimization process

init<-c(kout_pe_init,kout_ne_init,kout_in_ne_init)
mconstr<-rbind(diag(1,3,3),c(-1,0,0),c(0,-1,0),c(0,0,-1))

```

```

vconstr<-c(c(kout_pe_min,kout_ne_min,kout_in_ne_min),c(-kout_pe_max),c(-kout_ne_max),c(-kout_in_ne_max))

# Parameters of Model Bu are optimized by a two-stage minimization of the sum of squared errors,rss
# 1. pbase, kin_pe and kin_in_ne are computed by linear regression according to values for kout_pe,
# kout_ne and kout_in_ne
# 2. kout_pe,kout_ne and kout_in_ne are optimized with constrOptim function

rss<-function(param){
  kout_pe<-param[1]
  kout_ne<-param[2]
  kout_in_ne<-param[3]
  datareg<-list(pe=pe_1exp(kout_pe)[extract],ne=ne_2exp(kout_ne,kout_in_ne)[extract],p=pmes[extract])
  reg<-lm(p~pe+ne,datareg)
  pmod<-reg$coefficients[1]+reg$coefficients[2]*pe_1exp(kout_pe)+reg$coefficients[3]*ne_2exp(kout_ne,kout_in_ne)
  rss<-sum((pmes[extract]-pmod[extract])^2)
}

fit<-constrOptim(theta=init,f=rss,grad=NULL,ui=mconstr,ci=vconstr,control=list(maxit=3000))

kout_pe<-fit$par[1]
kout_ne<-fit$par[2]
kout_in_ne<-fit$par[3]
datareg<-list(pe=pe_1exp(kout_pe)[extract],ne=ne_2exp(kout_ne,kout_in_ne)[extract],p=pmes[extract])
reg<-lm(p~pe+ne,datareg)
pbase<-reg$coefficients[1]
kin_pe<-reg$coefficients[2]
kin_in_ne<-reg$coefficients[3]
pmod<-pbase+kin_pe*pe_1exp(kout_pe)-kin_in_ne*ne_2exp(kout_ne,kout_in_ne)
}

if (model == "TI"){

# Fixing starting values and restrictions of model parameters for optimization process

```

```

init<-c(kout_pe_init,kout_s_init,kinhib_init)
mconstr<-rbind(diag(1,3,3),c(-1,0,0),c(0,-1,0),c(0,0,-1))
vconstr<-c(c(kout_pe_min,kout_s_min,0),c(-kout_pe_max),c(-kout_s_max),c(-kinhib_max))

# Parameters of Model TI are optimized by a two-stage minimization of the sum of squared errors,rss
# 1. pbase and kin_pe are computed by linear regression according to values for kout_pe, kout_s and kinhib
# 2. kout_pe, kout_s and kinhib are optimized with constrOptim function

rss<-function(param){
  kout_pe<-param[1]
  kout_s<-param[2]
  kinhib<-param[3]
  datareg<-list(pe=pe_TI(kout_pe,kout_s,kinhib)[extract],p=pmes[extract])
  reg<-lm(p~pe,datareg)
  pmod<-reg$coefficients[1]+reg$coefficients[2]*pe_TI(kout_pe,kout_s,kinhib)
  rss<-sum((pmes[extract]-pmod[extract])^2)
}

fit<-constrOptim(theta=init,f=rss,grad=NULL,ui=mconstr,ci=vconstr,control=list(maxit=3000))

kout_pe<-fit$par[1]
kout_s<-fit$par[2]
kinhib<-fit$par[3]
datareg<-list(pe=pe_TI(kout_pe,kout_s,kinhib)[extract],p=pmes[extract])
reg<-lm(p~pe,datareg)
pbase<-reg$coefficients[1]
kin_s<-reg$coefficients[2]
pmod<-pbase+kin_s*pe_TI(kout_pe,kout_s,kinhib)
}

if (model == "TF"){

# Fixing starting values and restrictions of model parameters for optimization process

```

```

init<-c(kout_pe_init,kout_s_init,kout_ne_init)
mconstr<-rbind(diag(1,3,3),c(-1,0,0),c(0,-1,0),c(0,0,-1))
vconstr<-c(c(kout_pe_min,kout_s_min,kout_ne_min),c(-kout_pe_max),c(-kout_s_max),c(-kout_ne_max))

# Parameters of Model TF are optimized by a two-stage minimization of the sum of squared errors,rss
# 1. pbase, kin_pe and kin_ne are computed by linear regression according to values for kout_pe, kout_s and kout_ne
# 2. kout_pe, kout_s and kout_ne are optimized with constrOptim function

rss<-function(param){
  kout_pe<-param[1]
  kout_s<-param[2]
  kout_ne<-param[3]
  datareg<-list(pe=pe_T(kout_pe,kout_s)[extract],ne=ne_1exp(kout_ne)[extract],p=pmes[extract])
  reg<-lm(p~pe+ne,datareg)
  pmod<-reg$coefficients[1]+reg$coefficients[2]*pe_T(kout_pe,kout_s)+reg$coefficients[3]*ne_1exp(kout_ne)
  rss<-sum((pmes[extract]-pmod[extract])^2)
}

fit<-constrOptim(theta=init,f=rss,grad=NULL,ui=mconstr,ci=vconstr,control=list(maxit=3000))

kout_pe<-fit$par[1]
kout_s<-fit$par[2]
kout_ne<-fit$par[3]
datareg<-list(pe=pe_T(kout_pe,kout_s)[extract],ne=ne_1exp(kout_ne)[extract],p=pmes[extract])
reg<-lm(p~pe+ne,datareg)
pbase<-reg$coefficients[1]
kin_s<-reg$coefficients[2]
kin_ne<-reg$coefficients[3]
pmod<-pbase+kin_s*pe_T(kout_pe,kout_s)-kin_ne*ne_1exp(kout_ne)
}

if (model == "TF2"){

```

```

# Fixing starting values and restrictions of model parameters for optimization process

init<-c(kout_pe_init,kout_s_init,kout_ne_init,kout_in_ne_init)
mconstr<-rbind(diag(1,4,4),c(-1,0,0,0),c(0,-1,0,0),c(0,0,-1,0),c(0,0,0,-1))
vconstr<-c(c(kout_pe_min,kout_s_min,kout_ne_min,kout_in_ne_min),
           c(-kout_pe_max),c(-kout_s_max),c(-kout_ne_max),c(-kout_in_ne_max))

# Parameters of Model TF2 are optimized by a two-stage minimization of the sum of squared errors,rss
# 1. pbase, kin_pe and kin_in_ne are computed by linear regression according to values for kout_pe,
# kout_s, kout_ne and kout_in_ne
# 2. kout_pe, kout_s, kout_ne and kout_in_ne are optimized with constrOptim function

rss<-function(param){
  kout_pe<-param[1]
  kout_s<-param[2]
  kout_ne<-param[3]
  kout_in_ne<-param[4]
  datareg<-list(pe=pe_T(kout_pe,kout_s)[extract],ne=ne_2exp(kout_ne,kout_in_ne)[extract],p=pmes[extract])
  reg<-lm(p~pe+ne,datareg)
  pmod<-reg$coefficients[1]+reg$coefficients[2]*pe_T(kout_pe,kout_s)
  +reg$coefficients[3]*ne_2exp(kout_ne,kout_in_ne)
  rss<-sum((pmes[extract]-pmod[extract])^2)
}

fit<-constrOptim(theta=init,f=rss,grad=NULL,ui=mconstr,ci=vconstr,control=list(maxit=3000))

kout_pe<-fit$par[1]
kout_s<-fit$par[2]
kout_ne<-fit$par[3]
kout_in_ne<-fit$par[4]
datareg<-list(pe=pe_T(kout_pe,kout_s)[extract],ne=ne_2exp(kout_ne,kout_in_ne)[extract],p=pmes[extract])
reg<-lm(p~pe+ne,datareg)
pbase<-reg$coefficients[1]
kin_pe<-reg$coefficients[2]
kin_in_ne<-reg$coefficients[3]

```

```

pmod<-pbase+kin_pe*pe_T(kout_pe,kout_s)-kin_in_ne*ne_2exp(kout_ne,kout_in_ne)
}

if (model == "TIF"){

# Fixing starting values and restrictions of model parameters for optimization process

init<-c(kout_pe_init,kout_s_init,kinhib_init,kout_ne_init)
mconstr<-rbind(diag(1,4,4),c(-1,0,0,0),c(0,-1,0,0),c(0,0,-1,0),c(0,0,0,-1))
vconstr<-c(c(kout_pe_min,kout_s_min,0,kout_ne_min),
           c(-kout_pe_max),c(-kout_s_max),c(-kinhib_max),c(-kout_ne_max))

# Parameters of Model TIF are optimized by a two-stage minimization of the sum of squared errors,rss
# 1. pbase, kin_pe and kin_ne are computed by linear regression according to values for kout_pe,
# kout_s, kout_ne and kinhib
# 2. kout_pe, kout_s, kout_ne and kinhib are optimized with constrOptim function

rss<-function(param){
  kout_pe<-param[1]
  kout_s<-param[2]
  kinhib<-param[3]
  kout_ne<-param[4]
  datareg<-list(pe=pe_TI(kout_pe,kout_s,kinhib)[extract],ne=ne_1exp(kout_ne)[extract],p=pmes[extract])
  reg<-lm(p~pe+ne,datareg)
  pmod<-reg$coefficients[1]+reg$coefficients[2]*pe_TI(kout_pe,kout_s,kinhib)
  +reg$coefficients[3]*ne_1exp(kout_ne)
  rss<-sum((pmes[extract]-pmod[extract])^2)
}

fit<-constrOptim(theta=init,f=rss,grad=NULL,ui=mconstr,ci=vconstr,control=list(maxit=3000))

kout_pe<-fit$par[1]
kout_s<-fit$par[2]
kinhib<-fit$par[3]

```



```

kout_ne<-fit$par[4]
datareg<-list(pe=pe_TI(kout_pe,kout_s,kinhib)[extract],ne=ne_1exp(kout_ne)[extract],p=pmes[extract])
reg<-lm(p~pe+ne,datareg)
pbase<-reg$coefficients[1]
kin_pe<-reg$coefficients[2]
kin_ne<-reg$coefficients[3]
pmod<-pbase+kin_pe*pe_TI(kout_pe,kout_s,kinhib)-kin_ne*ne_1exp(kout_ne)
}

# Computation of measures of the goodness-of-fit

sct<-sum((pmes[extract]-mean(pmes[extract]))^2)
scr<-sum((pmes[extract]-pmod[extract])^2)
r2<-1-(scr/sct)
r2adj<-1-(n-1)/(n-p)*(1-r2)
se<-sqrt(scr/(n-p))
logL<-0.5*n*(log(2*pi)+1-log(n)+log(sum((pmes[extract]-pmod[extract])^2)))
aic<-2*p-2*logL
aicc<-aic+2*p*(p+1)/(n-p-1)

Fit_Model<-list(r2=r2,r2adj=r2adj,se=se,logL=logL,aic=aic,aicc=aicc,converg=fit$convergence,
               pbase=pbase,kout_pe=kout_pe,kout_ne=kout_ne,kout_in_ne=kout_in_ne,kout_s=kout_s,
               kinhib=kinhib,kin_pe=kin_pe,kin_ne=kin_ne,kin_in_ne=kin_in_ne,kin_s=kin_s)

Fit_Model<-Fit_Model
}

```

**Supplemental document 3.** Alternative solution of the models and impact on model estimates

**Solution # 1**

It is the result obtained with the procedure used in the main study given in the supplemental document 2. Model parameters were fitted within the region defined by the following constraints:

- kout<sub>pe</sub> in [0.01; 0.07] and starting point = 0.03 for all models
- kout<sub>ne</sub> in [0.06; 1] and starting point = 0.1 for models Ba, Bu, TF, TF2 and TIF
- kout<sub>in\_ne</sub> in [0.1; 2] and starting point = 0.5 for models Bu and TF2
- kout<sub>s</sub> in [0.05; 1] and starting point = 0.1 for models TI, TF, TF2 and TIF
- k<sub>in\_i</sub> in [0; 0.01] and starting point = 0.001 for models TI and TIF

**Solution # 2**

Alternative procedure using upper bound for kout<sub>pe</sub> lower than in the main study:

- kout<sub>pe</sub> in [0.01; 0.03] and starting point = 0.02 for all models

Table 1 gives RSS minimized using the two procedures. Small differences were observed for each model.

**Table 1 : Mean ± SD of residual sum of squares (RSS) for solution #1 and solution #2.**

Model	RSS #1	RSS #2	Difference #1 - #2
Ba	477 ± 272	479 ± 261	-2 ± 28
Bu	465 ± 250	462 ± 242	3 ± 25
TI	562 ± 288	568 ± 287	-6 ± 10
TF	475 ± 271	479 ± 261	-4 ± 29
TF2	463 ± 256	460 ± 242	3 ± 27
TIF	468 ± 272	481 ± 265	-12 ± 27

No significant difference between solutions #1 and #2.

Tables 2 and 3 give the estimates of parameters of Model Ba and Model Bu showing large differences between the 2 solutions.

**Table 2 : Mean  $\pm$  SD of estimates for Model Ba.**

	<b>Solution #1</b>	<b>Solution #2</b>
$P_{\text{base}}$ (% of pre-training)	101.3 $\pm$ 0.7	101.1 $\pm$ 0.9
$k_{\text{out}_{\text{pe}}}$ (day <sup>-1</sup> )	0.0605 $\pm$ 0.0192	0.0209 $\pm$ 0.0104 *
$\tau_{\text{pe}}$ (day)	23.4 $\pm$ 25.6	63.6 $\pm$ 34.8 *
$k_{\text{in}_{\text{pe}}}$ (a.u.)	0.0283 $\pm$ 0.0153	0.0043 $\pm$ 0.0022 *
$k_{\text{out}_{\text{ne}}}$ (day <sup>-1</sup> )	0.157 $\pm$ 0.031	0.237 $\pm$ 0.048 *
$\tau_{\text{ne}}$ (day)	6.57 $\pm$ 1.04	4.36 $\pm$ 0.77 *
$k_{\text{in}_{\text{ne}}}$ (a.u.)	0.0666 $\pm$ 0.0260	0.0425 $\pm$ 0.0128 *

a.u. : arbitrary unit.

\* : significant difference between solutions #1 and #2.

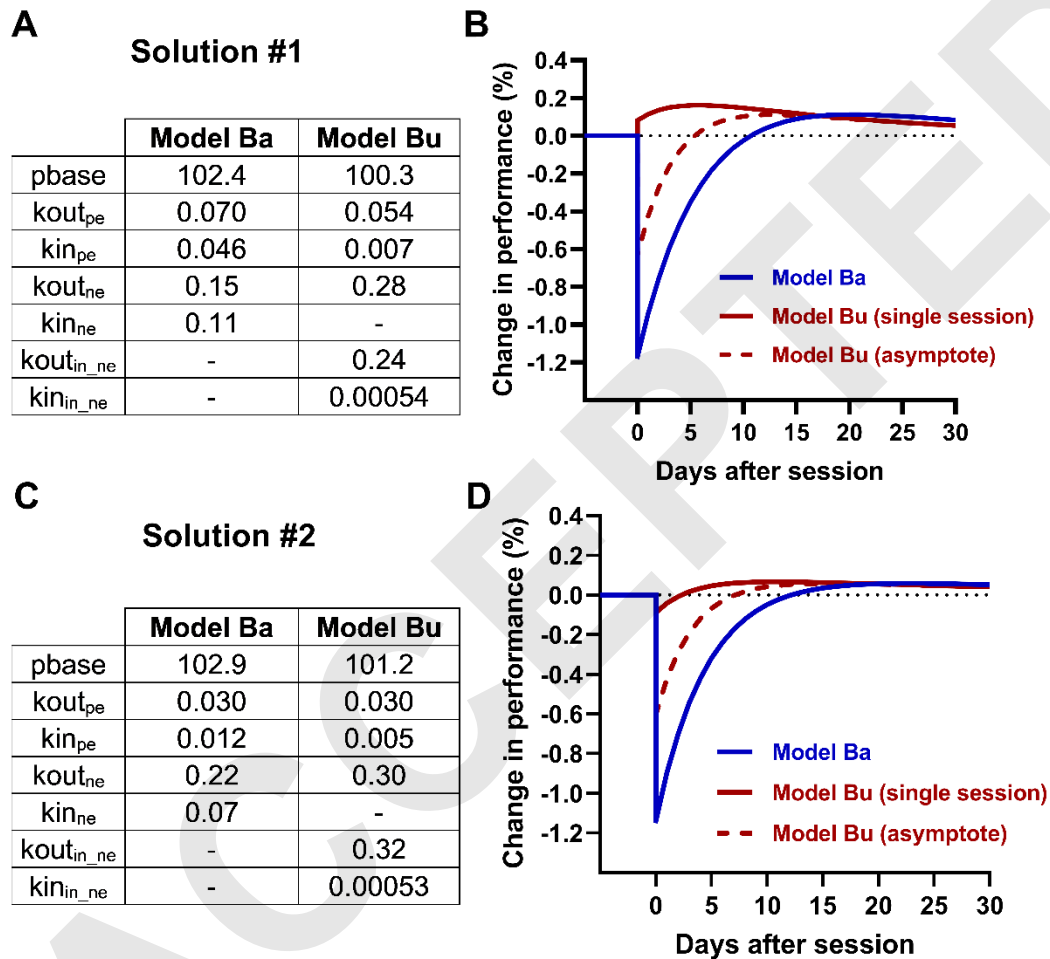
**Table 3 : Mean  $\pm$  SD of estimates for Model Bu.**

	<b>Solution #1</b>	<b>Solution #2</b>
$P_{\text{base}}$ (%)	99.4 $\pm$ 0.7	99.9 $\pm$ 0.7 *
$k_{\text{out}_{\text{pe}}}$ (day <sup>-1</sup> )	0.0455 $\pm$ 0.0207	0.0227 $\pm$ 0.0101 *
$\tau_{\text{pe}}$ (day)	34.3 $\pm$ 32.7	57.6 $\pm$ 33.6 *
$k_{\text{in}_{\text{pe}}}$ (a.u.)	0.0114 $\pm$ 0.0074	0.0038 $\pm$ 0.0014 *
$k_{\text{out}_{\text{ne}}}$ (day <sup>-1</sup> )	0.306 $\pm$ 0.067	0.351 $\pm$ 0.059 *
$\tau_{\text{ne}}$ (day)	3.39 $\pm$ 0.67	2.93 $\pm$ 0.48 *
$k_{\text{out}_{\text{in}_{\text{ne}}}}$ (day <sup>-1</sup> )	0.57 $\pm$ 0.56	1.20 $\pm$ 0.69 *
$\tau_{\text{in}_{\text{ne}}}$ (day)	3.75 $\pm$ 3.36	1.32 $\pm$ 1.09 *
$k_{\text{in}_{\text{in}_{\text{ne}}}}$ (a.u.)	0.000511 $\pm$ 0.000263	0.000650 $\pm$ 0.000216

a.u. : arbitrary unit.

\* : significant difference between solutions #1 and #2.

Figure 1 shows the impulse response estimated using the two solutions for median TL for one swimmer. Tables 4 and 5 compare the characteristics of impulse responses for the entire group of swimmers. Despite the large difference in the estimates of parameters, the impulse responses for Model Ba, Model Bu for single bout and last bout effects at the asymptote were close when using the two set of parameters. Only  $p_g$  showed large difference between the two solutions.



**Figure 1 : Impulse response to median TL with the estimates from Model Ba and Model Bu found for one swimmer. Panel A and B : estimates of model parameters and impulse responses for solution #1. Panel C and D : estimates of model parameters and impulse responses for solution #2.**

**Table 4 : Mean  $\pm$  SD of characteristics of impulse response to median TL for Model Ba.**

	<b>Solution #1</b>	<b>Solution #2</b>
$\Delta p_0$ (% of pre-training)	0.92 $\pm$ 0.25	0.92 $\pm$ 0.24
$t_n$ (day)	10.5 $\pm$ 1.14	11.2 $\pm$ 1.34 *
$p_g$ (% of pre-training)	0.096 $\pm$ 0.027	0.054 $\pm$ 0.014 *
$t_g$ (day)	20.8 $\pm$ 2.7	23.0 $\pm$ 2.5 *

\* : significant difference between solutions #1 and #2.

**Table 5 : Mean  $\pm$  SD of characteristics of impulse response to median TL for Model Bu given for the response to an isolated training bout (single session) and to the last bout at steady state after daily repetitions (Asymptote).**

	<b>Solution #1</b>	<b>Solution #2</b>
<i>Single session</i>		
$\Delta p_0$ (% of pre-training)	0.03 $\pm$ 0.31	0.31 $\pm$ 0.18 *
$t_n$ (day)	1.4 $\pm$ 2.2	4.3 $\pm$ 1.7 *
$p_g$ (% of pre-training)	0.17 $\pm$ 0.12	0.09 $\pm$ 0.03 *
$t_g$ (day)	7.6 $\pm$ 4.4	13.0 $\pm$ 2.8 *
<i>Asymptote</i>		
$\Delta p_0$ (% of pre-training)	0.62 $\pm$ 0.130	0.55 $\pm$ 0.12
$t_n$ (day)	5.2 $\pm$ 0.9	6.1 $\pm$ 1.2
$p_g$ (% of pre-training)	0.11 $\pm$ 0.05	0.06 $\pm$ 0.01 *
$t_g$ (day)	12.9 $\pm$ 1.8	14.9 $\pm$ 2.7 *

\* : significant difference between procedure #1 and procedure #2.

Table 6 compares the characteristics of optimal taper computed using the two set of parameters for Model Bu. No statistical difference was observed between the two solutions with the exception of optimal duration with step reduction of training but the difference was lower than 1 day.

**Table 6 : Mean  $\pm$  SD of estimated characteristics of the optimal simulated taper according to the form of the training reduction after regular training (median TL) and a step increase of 20% from regular training for 28 days.**

	Form of reduction of training during taper		
	Step	Linear	Exponential
<i>Regular training</i>			
Reduction (%)			
Solution #1	87 $\pm$ 16	67 $\pm$ 13	62 $\pm$ 13
Solution #2	87 $\pm$ 7	73 $\pm$ 8	65 $\pm$ 9
Duration (days)			
Solution #1	6.5 $\pm$ 1.2	9.0 $\pm$ 1.1	8.2 $\pm$ 0.6
Solution #2	7.1 $\pm$ 1.2 *	8.8 $\pm$ 0.6	8.6 $\pm$ 1.0
Highest perf. (%)			
Solution #1	103.2 $\pm$ 1.1	103.2 $\pm$ 1.1	103.2 $\pm$ 1.1
Solution #2	103.6 $\pm$ 1.3	103.6 $\pm$ 1.3	103.6 $\pm$ 1.3
<i>Regular training +20%</i>			
Reduction (%)			
Solution #1	86 $\pm$ 14	71 $\pm$ 11	66 $\pm$ 13
Solution #2	88 $\pm$ 8	74 $\pm$ 6	66 $\pm$ 7
Duration (days)			
Solution #1	7.6 $\pm$ 0.7	9.5 $\pm$ 1.7	8.7 $\pm$ 1.0
Solution #2	8.1 $\pm$ 1.0 *	9.5 $\pm$ 1.0	9.3 $\pm$ 1.1
Highest perf. (%)			
Solution #1	103.9 $\pm$ 1.3	103.8 $\pm$ 1.3	103.8 $\pm$ 1.3
Solution #2	104.3 $\pm$ 1.6	104.3 $\pm$ 1.6	104.3 $\pm$ 1.6

\* : significant difference between solutions #1 and #2.

Table 7 compares the accuracy of performance prediction computed using the two set of parameters for Model Ba and Bu. Statistical difference between the two solutions was observed for MAPE for the two validation datasets. These small differences lower than 0.1 % are explained by significant difference in systematic error : greater for solution #2 than solution #1 for Validation dataset 1 and lower for solution #2 than solution #1 for Validation dataset 2.

**Table 7 : Accuracy of the performance prediction for validation dataset 1 and 2 using cycles #1 and #2 for training models**

Data set	Model Ba		Model Bu	
	Solution #1	Solution #2	Solution #1	Solution #2
<i>Validation 1</i>				
Systematic error	0.16 ± 0.40	0.49 ± 0.43 *	0.21 ± 0.42	0.47 ± 0.50 *
Typical error	2.40 ± 0.74	2.44 ± 0.75 *	2.56 ± 0.80	2.54 ± 0.80
MAPE	2.02 ± 0.65	2.12 ± 0.66 *	2.17 ± 0.70	2.22 ± 0.70 *
<i>Validation 2</i>				
Systematic error	-1.74 ± 1.23	-1.48 ± 1.15 *	-1.64 ± 1.15	-1.50 ± 1.16 *
Typical error	2.43 ± 0.51	2.51 ± 0.59	2.32 ± 0.49	2.36 ± 0.50
MAPE	2.69 ± 1.23	2.60 ± 0.85 *	2.56 ± 0.79	2.52 ± 0.78

Systematic and typical error are expressed as a percentage of first performance of Season 1.

\* : significant difference between solutions #1 and #2.