

**UNIVERSIDADE DE TRÁS-OS-MONTES E ALTO DOURO**

**LONGITUDINAL MODELING OF YOUNG SWIMMERS'  
PERFORMANCE AND BIOMECHANICS:  
IDENTIFICATION, DEVELOPMENT AND FOLLOW-UP**

**DOUTORAMENTO EM CIÊNCIAS DO DESPORTO**

**JORGE FILIPE ESTRELA MORAIS**

**ORIENTADORES:** TIAGO MANUEL CABRAL DOS SANTOS BARBOSA

ANTÓNIO JOSÉ ROCHA MARTINS DA SILVA

DANIEL ALMEIDA MARINHO



**VILA REAL 2016**



This academic thesis was submitted with the aim of obtaining a doctoral degree in Sport Sciences according to the provision of the Portuguese Decree-Law 107/2008, June 25<sup>th</sup>.



## Funding

The Portuguese Science and Technology Foundation (FCT) funded this thesis, under the Human Potential Operating Program, supported by the European Social Found (ESF), grant SFRH/BD/76287/2011.





To my parents...  
Dedicated to my beloved mother...



## Table of Contents

Acknowledgments .....	xi
List of publications .....	xiii
Figures Index.....	xv
Tables Index.....	xvii
List of Abbreviations .....	xix
Abstract .....	xxi
Resumo .....	xxiii
General Introduction .....	1
Chapter 1 .....	3
Chapter 2.....	29
Chapter 3.....	51
Chapter 4.....	75
General Discussion and Conclusions .....	97
General References .....	103



## Acknowledgments

The development of this research, culminated in this doctoral thesis would not be possible without the contribution of so many people, therefore my sincere appreciation and gratitude.

First, to my supervisor Tiago Manuel Cabral dos Santos Barbosa (PhD) for all the guidance, support and motivation since the beginning of my academic journey, and most of all for his friendship.

To my co-supervisors, António José Rocha Martins da Silva (PhD) and Daniel Almeida Marinho (PhD) for their disposal in the manuscript preparation and setting the logistics for the data collection.

To those who participated in the data collection and/or in the manuscript preparation: Eduardo Pires, Henrique Neiva (PhD), Marc Moreira, Mário Costa (PhD), Mário Marques (PhD), Nuno Garrido (PhD), Pedro Forte, Sérgio Jesus, Tiago Poço, Vasco Lopes, Vítor Lopes (PhD).

To all athletes, coaches and researchers involved in this research project, showing full cooperation and disposal.



## List of publications

### The following parts of this thesis are published:

Morais JE, Jesus S, Lopes V, Garrido ND, Silva AJ, Marinho DA, Barbosa TM. (2012). Linking selected kinematical, anthropometric and hydrodynamic variables to young swimmer performance. *Pediatr Exerc Sci.* 24: 649-664.

Morais JE, Marques MC, Marinho DA, Silva AJ, Barbosa TM. Longitudinal modeling in sports: young swimmers' performance and biomechanics profile. (2014). *Hum Mov Sci.* 37: 111-122.

Morais JE, Silva AJ, Marinho DA, Seifert L, Barbosa TM. (2015). Cluster stability as a new method to assess changes in performance and its determinant factors over a season in young swimmers. *Int J Sports Physiol Perf.* 10: 261-268.

### The following part of the present thesis is submitted and under review:

Morais JE, Silva AJ, Marinho DA, Lopes VP, Barbosa TM. (Submitted). Determinant factors of long-term performance development in young swimmers. *Int J Sports Physiol Perf.*



## Figures Index

Chapter 1	Figure 1. Theoretical path-flow model.	10
Chapter 1	Figure 2. Overall confirmatory path-flow model including all variables computed (2A) and deleting variable that allowed to reduce the residual error and improve the goodness-of-fit (2B) with the subsequent recomputation of remain data.	18
Chapter 2	Figure 1. Training volume per week (in km) throughout the season. # – evaluation moments (M <sub>i</sub> ).	36
Chapter 2	Figure 2. Study design scheme. M – moment; Wk – week; # – week's number.	36
Chapter 2	Figure 3. Theoretical model.	37
Chapter 2	Figure 4. Growth confirmatory models for performance and effects of the selected variables.	43
Chapter 3	Figure 1. Territorial map of the two canonical discriminant functions in moment 1 (M1-A), moment 2 (M2-B) and moment 3 (M3-C), respectively.	64
Chapter 4	Figure 1. Training volume per week (in km) in each season, and the performance variation.	81
Chapter 4	Figure 2. The timeline for the data collection over the three seasons (10 evaluation moments).	82



## Tables Index

Chapter 1	Table 1. Overall, boys and girls descriptive statistics of anthropometric, kinematic, hydrodynamic and swimming performance variables.	16
Chapter 1	Table 2. Overall, boys and girls Person's correlation coefficients between swimming performance and remain variables selected.	17
Chapter 2	Table 1. Descriptive statistics for selected kinematic and hydrodynamic variables in each evaluation moment.	41
Chapter 3	Table 1. Performance and its determining factors for swimmers classified in each cluster assessed at baseline (moment 1), mid-season (moment 2), and end-season (moment 3).	63
Chapter 3	Table 2. Number of swimmers re-classified in each cluster between baseline (moment 1) and mid-season (moment 2), between mid-season (moment 2) and end-season (moment 3), and between baseline (moment 1) and end-season (moment 3). Distances between cluster centers for each pairwise comparison of clusters at each moment are also shown.	66
Chapter 4	Table 1. Descriptive statistics and variation (%; 95% CI) of the anthropometrics between season #1 – early and season #3 – late.	86
Chapter 4	Table 2. Descriptive statistics and variation (%; 95% CI) of the technical and performance data between season #1 – early and season #3 – late.	87
Chapter 4	Table 3. Parameters of the two models computed with standard errors (SE) and 95% confidence intervals (CI).	89



## List of Abbreviations

AS – arm span

a.u. – arbitrary unit

$\beta$  – beta value for standardized coefficients

$D_a$  – active drag

$C_{D_a}$  – coefficient of active drag

CFI – comparative fit index

CI – confidence interval

CP – chest perimeter

dv – intra-cyclic velocity fluctuation

$\eta^2$  – total eta square

FINA – Fédération Internationale de Natation

HSA – hand surface area

ICC – intra-class correlation coefficient

ICEPT – intercept

M – evaluation moment

n – number of subjects

% – percentage

$\eta_p$  – propelling efficiency

$P_d$  – power to overcome drag

Perf – swimming performance

r – Pearson product-moment correlation coefficient

$r^2$  – coefficient of determination

SD – standard deviation

SCM – short course meter (swimming pool)

SF – stroke frequency

SI – stroke index

SL – stroke length

v – swimming velocity

VAR – variable

$\chi^2/df$  – chi-square/degrees of freedom

z – standardized z-scores

## Abstract

These days, talent identification programs are becoming extremely useful providing important details on the performance determinants in young swimmers and how it change over time. This information might help swimmers to excel, and eventually reach an elite level. Therefore, the main aims of this thesis were: (i) to identify the main determinants of the young swimmers' performance (study #1); (ii) observe and understand the young swimmers' performance changes, and the determinant factors associated to it over one season (study #2); (iii) identify, classify and follow-up young swimmers, based on their performance and determinant factors, as well as their stability over one season (study #3); (iv) and develop a performance predictor model, over three consecutive seasons, based on the swimmers biomechanical profile (study #4). In study #1 it was computed a structural equation model for young swimmers' performance based on selected kinematic, anthropometric, efficiency and hydrodynamic variables. It was verified that young swimmers' performance depends from a set of anthropometric, kinematic, efficiency and hydrodynamic factors. In study #2 a latent growth curve model was developed. Young swimmers' performance significantly improved, and with a significant inter-variability. Different determinant factors were responsible for such improvement in each evaluation moment. In study #3 a cluster analysis was used to classify, identify and follow-up the performance and its determinant factors. It was showed that within an age-group of prepubescent swimmers, three sub-groups with similar biomechanical characteristics were found. In study #4, a predictive model, over three consecutive seasons was developed based on biomechanics. The predictive model included an anthropometric, a kinematic and an efficiency factor, showing the multifactorial phenomenon that swimming is. The main conclusions of this thesis were that anthropometrics, kinematics, efficiency and hydrodynamics characterize young swimmers' profile and their performance, showing an improvement during the time-frames evaluated.

**Key-words:** *talent ID, training, kinematics, kinetics, anthropometrics, performance, modeling*



## Resumo

Os programas de identificação de talentos estão a tornar-se de extrema utilidade, fornecendo dados importantes sobre os determinantes da performance em nadadores jovens, e como esta evolui ao longo do tempo. Esta informação pode ajudá-los a atingirem um nível de elite. No entanto, a literatura baseia-se estudos transversais ou em estudos longitudinais de curto prazo. Os principais objetivos desta tese foram: (i) identificar os principais determinantes da performance em nadadores jovens (estudo #1); (ii) observar e entender a evolução da performance em nadadores jovens, e os fatores determinantes associados, durante uma época (estudo #2); (iii) identificar, classificar e acompanhar nadadores jovens, com base na sua performance e fatores determinantes, bem como a sua estabilidade durante uma época (estudo #3); (iv) e desenvolver um modelo preditor da performance, durante três épocas consecutivas com base no perfil biomecânico dos nadadores (estudo #4). No estudo #1 desenvolveu-se um modelo de equações estruturais para a performance, com base em variáveis cinemáticas, antropométricas, eficiência e hidrodinâmicas. Verificou-se que a performance depende de um conjunto de fatores antropométricos, cinemáticos, eficiência e hidrodinâmicos. No estudo #2 foi desenvolvido um modelo de crescimento latente. A performance melhorou significativamente, e com uma inter-variabilidade significativa. Diferentes fatores determinantes foram responsáveis por essa melhoria em cada um dos momentos de avaliação. No estudo #3 foi utilizada a análise de clusters para classificar, identificar e acompanhar a performance e os seus fatores determinantes. Verificou-se em nadadores pré-púberes, existirem três subgrupos com características biomecânicas semelhantes. No estudo #4, foi desenvolvido um modelo preditivo, durante três épocas consecutivas, com base em fatores biomecânicos. O modelo preditivo incluiu uma variável antropométrica, uma cinemática e uma de eficiência, evidenciando que a natação competitiva é um fenómeno multifatorial. As principais conclusões foram que variáveis antropométricas, cinemáticas, eficiência e hidrodinâmicas caracterizaram o perfil dos nadadores jovens e a sua performance, mostrando uma melhoria nos períodos de tempo avaliados.

**Palavras-chave:** *identificação de talentos, treino, cinemática, cinética, antropometria, performance, modelação*



## General Introduction

Research on young swimmers' performance and its determinants are scarce and mainly based on cross-sectional designs (Saavedra et al., 2010; Jurimae et al., 2007; Geladas et al., 2005). For one side, this research design provides some information on the main determinants of young swimmers' performance. Literature reports biomechanics (being anthropometrics, kinematics and hydrodynamics under this scientific field) as the main responsible for the young swimmers' performance (Barbosa et al., 2014). On the other hand, these research designs unable to have an understanding of the swimmers' changes over time, as well as, changes in the interaction among different determinants.

The best practice to gather a comprehensive insight on the relationships between all the factors that have an effect on swimming performance is designing longitudinal researches (Barbosa et al., 2015; Garrido et al., 2010; Latt et al., 2009a,b), albeit the number of papers reporting this are rather scarce. Having that said, as far as longitudinal studies reported in the literature goes, there are a few concerns that should be highlighted: (i) the sample: small and underpowered (on top of that, the subjects recruited are not always talented swimmers); (ii) the absence of modeling procedures and/or the data analysis selected not being the most cutting-edge and insightful; (iii) the time-frame (i.e. mostly short time-frames and/or with few evaluation moments) and; (iv) the relationship between the external and internal training loads are not provided or at least reported in a comprehensive way.

One of the new trends in sports sciences is the identification and development of talented young athletes. Overall, this process includes the identification, characterization and follow-up of young talented athletes (i.e. performance and its determinants) (Erlandson et al., 2008; Matthys et al., 2013; Robertson et al., 2014). In competitive swimming, as national and world records keep being broken, practitioners and researchers are willing to anticipate who will be the next top-ranked swimmer. Therefore, as in other sports, swimming fraternity is also keen to have a deep insight on this process.

Until recently, classical research designs and data analysis procedures (e.g. analysis of variance and regression models) selected in sports performance were not helpful

in gathering insight about such highly dynamic and complex relationships. Evidence has been gathered lately on this in adult/elite swimmers (Komar et al., 2014; Costa et al., 2013) despite definitive answers are needed. Yet, little or almost nothing is known about such relationships in young swimmers.

The majority of studies about young swimmers are based on correlations, analysis of variance and regression models (Vitor and Bohme, 2012; Latt et al., 2009a,b; Jurimae et al., 2007). These studies only provide the magnitude of association between the performance and its determinant factors, unable to report how those determinants interplay. One might consider that, as their adult/elite counterparts, young swimmers' performance depends from interaction between several variables, belonging to different scientific fields (Barbosa et al., 2010a). That said, even though the same determinants can or cannot be in play, the partial contribution to the main outcome (i.e. performance) might be different from what has been reported earlier for adult/elite swimmers. Being biomechanics the domain with higher association to the young swimmers' performance, one can speculate modeling such relationships would be interesting. Data on the change of the performance over time is also scarce. There is no information available in the literature if the partial contribution of each determinant is kept the same over time or not and if so, is there any relationship between these variations and the external training load?

Therefore, the aim of this thesis was to identify and follow-up young swimmers' performance and its determinant factors over time.

The thesis features four research studies that enables to breakdown the main aim into:

- identifying the main determinants of the young swimmers' performance (chapter 1);
- assess the changes in young swimmers' performance and its determinants over one season (chapter 2);
- identify, classify and follow-up over one season young swimmers, based on their performance and determinant factors, as well as assess their stability (chapter 3);
- predict the performance based on the changes in the swimmers' biomechanics over three consecutive seasons (chapter 4).

---

## **Chapter 1**

**Linking selected kinematic, anthropometric and hydrodynamic variables to young swimmer performance**

---



**Abstract**

The aim of this study was to develop a structural equation model (i.e. a confirmatory technique that analyzes relationships among observed variables) for young swimmers' performance based on selected kinematic, anthropometric and hydrodynamic variables. A total of 114 subjects (73 boys and 41 girls of mean age of  $12.31 \pm 1.09$  years;  $47.91 \pm 10.81$  kg body mass;  $156.57 \pm 10.90$  cm height and Tanner stages 1-2) were evaluated. The variables assessed were the: (i) 100-m freestyle performance; (ii) stroke index; (iii) intra-cyclic velocity fluctuation; (iv) stroke length; (v) active drag; (vi) arm span and; (vii) hand surface area. All paths were significant ( $P < 0.05$ ). However, in deleting the path between the hand surface area and the stroke index, the model goodness-of-fit significantly improved. Swimming performance in young swimmers appeared to be dependent on swimming efficiency (i.e. stroke index), which is determined by the remaining variables assessed, except for the hand surface area. Therefore, young swimmer coaches and practitioners should design training programs with a focus on technical training enhancement (i.e. improving swimming efficiency).

**Key-words:** *confirmatory assessment, technical training, swim efficiency, predictive factors, swimming performance*



## Introduction

Swimming performance results from a multifactorial process that involves several scientific domains, such as the anthropometrics (Latt et al., 2009a; Geladas et al., 2005; Duché et al., 1993), hydrodynamics (Marinho et al., 2010a; Kjendlie and Stallman, 2008), kinematics (Barbosa et al., 2010a; Jurimae et al., 2007) and energetics (Greco et al., 2007; Poujade et al., 2002; Denadai et al., 2000). As in adult/elite swimmers, one of the main goals of swimming research is to identify the scientific domains and/or variables that predict swimming performance in children (i.e. young athletes) thereby enhancing the detection of future talent (Hohmann and Seidel, 2010; Silva et al., 2000). Nevertheless, research in young athletes ought to be less invasive, expensive and time-consuming than in adult/elite counterparts (Garrido et al., 2010). In this sense, several authors (Latt et al., 2010; Barbosa et al., 2010a; Kjendlie and Stallman, 2008) have estimated and/or measured variables in various scientific domains (i.e. anthropometric, hydrodynamic, kinematic and energetic) that are easy to collect and may predict performance and/or detect talented swimmers. Since swimming competition starts at an early age, it is important to know when and how these variables interact with each other, as well as with performance. Several authors studied these relationships (Barbosa et al., 2010a; Saavedra et al., 2010; Kjendlie and Stallman, 2008) aiming to describe and/or better understand this phenomenon. It is reported that young swimmers' performance is strongly related with anthropometric and kinematic variables (Barbosa et al., 2010a; Vitor and Böhme, 2010). Moreover, both sets of variables are affected by the processes of growth and maturation (Latt et al., 2009b). Given this rationale, it seems that kinematic variables are those that best explain young swimmers' performance. Swimming velocity ( $r^2 = -0.93$ ) and stroke frequency ( $r^2 = -0.78$ ) were highly correlated with 100-m freestyle performance (Latt et al., 2010). However, during growth and maturation processes, anthropometric variables are also related with swimming performance in young athletes (Saavedra et al., 2010; Latt et al., 2009a,b). The arm span (AS), seems to be a major performance determinant since it is correlated with stroke mechanics, namely the stroke length (SL) and stroke index (SI; Jurimae et al., 2007). Arm span ( $r^2 = 0.48$ ) and SI ( $r^2 = 0.78$ ) were reported as the best overall predictors in 100-m freestyle event (Latt et al., 2010). Moreover, hydrodynamic variables also play an important role in swimming performance (Vilas-

Boas et al., 2010) and are also commonly reported in studies involving young swimmers (Barbosa et al., 2010b; Marinho et al., 2010a). Understanding the relationships between human morphology and hydrodynamic resistance allows coaches to modify stroke mechanics to enhance performance (Benjanuvatra et al., 2001). Furthermore, active drag ( $D_a$ ) has an important role in swimming performance, being highly dependent on swimming technique (Kjendlie and Stallman, 2008).

A key question is to understand how these different scientific domains and variables interact to enhance swimming performance. In confirmatory research, analysis is driven by theoretical relationships among variables that are hypothesized and tested by the researchers. The present study aimed to confirm whether the hypothesized interaction takes place. A confirmatory model of such relationships based on existing exploratory research reported in the main literature could be useful, not only to prescribe appropriate periodization programs and training sets for young swimmers, but also to promote feasible and effective programs to detect and to select talent in competitive swimming. Structural equation modeling is a confirmatory technique (i.e. data analysis procedure) that assesses relationships among observed variables with the main goal of providing a quantitative test of the theoretical model hypothesized by the researchers. To our knowledge only one study has so far attempted to confirm correlates between young swimmers' performance and at least some of these scientific domains (Barbosa et al., 2010a). The present paper is a follow-up from that study but more specially focused on understanding and developing the biomechanical factor (i.e. quantifying the partial contribution of the biomechanics domain to young swimmers' performance) in the model reported by these authors (Barbosa et al., 2010a).

The aim of this study was to develop a structural equation model for performance in young swimmers based on selected kinematic, anthropometric and hydrodynamic variables. It was hypothesized that swimming performance in young swimmers might be related with these variables. The swimming performance is mainly related to swimming efficiency and this one to several kinematic, anthropometric and hydrodynamic variables.

## Methods

### Participants

A total of 114 young swimmers participating on a regular basis in regional and national level competitions volunteered as subjects. They comprised 73 boys and 41 girls with a chronological age of  $12.31 \pm 1.09$  years (overall:  $47.91 \pm 10.81$  kg of body mass;  $156.57 \pm 10.90$  cm of height and Tanner stages 1-2 assessed by self-evaluation; boys:  $12.72 \pm 1.03$  years old;  $47.41 \pm 10.09$  kg of body mass,  $157.20 \pm 11.17$  cm of height and Tanner stages 1-2 assessed by self-evaluation; girls:  $11.47 \pm 0.66$  years old;  $45.79 \pm 6.66$  kg of body mass,  $154.56 \pm 8.26$  cm of height and Tanner stages 1-2 by self-evaluation).

Coaches and parents gave their consent for swimmers' participation in this study and all procedures were in accordance to the Helsinki Declaration concerning human research. The Institutional Review Board of the University approved the study design.

### Study design

#### Theoretical model

The theoretical model was designed according to exploratory state-of-the-art research and to test it was the object of our research. Figure 1 presents the theoretical model adopted for swimming performance based on selected kinematic, anthropometric and hydrodynamic variables in young swimmers. Swimming performance is related to kinematic (Barbosa et al., 2010a), anthropometric (Latt et al., 2009a; Duché et al., 1993) and hydrodynamic (Kjendlie and Stallman, 2008) variables. It was suggested that swimming performance depends on the relationship between the swimmer morphology, hydrodynamic resistance and swimming stroke mechanics (Benjanuvatra et al., 2001). The sequence of the theoretical model was designed according to these facts. For anthropometric assessment the surface area of the dominant hand (HSA) was computed. It is known that the propulsive surface is a key variable in increasing propulsive forces (e.g. propulsive drag and lift force). However, to the best of our knowledge, there are no studies deploying this variable in young swimmers. The AS is a variable reported on a regular basis in talent detection and selection (Hohmann and Seidel, 2010; Silva et al., 2000). The AS strongly affects not only the SL but also some hydrodynamic variables related to the body

length (Kjendlie and Stallman, 2008). Swimming with lower drag at constant velocity reduces the energy cost of swimming (Marinho et al., 2010a). The hydrodynamic variable assessed was the  $D_a$ . The kinematic variables analyzed were the intra-cyclic velocity fluctuation ( $dv$ ; Barbosa et al., 2008), the SL (Craig and Pendergast, 1979) and the SI (Costill et al., 1985). Intra-cyclic velocity fluctuation is the result of the propulsive and drag forces that interact on the swimmer and thus allows an overall assessment of the stroke mechanics (Barbosa et al., 2008). Stability or minimal change in SL at a high value is associated with higher performances (Sidney et al., 2010). The SI is strongly related to the energy cost of swimming (Costill et al., 1985). Indeed, the SI is the swimming economy estimator most often cited by the scientific community. It describes the swimmers' ability to move at a given velocity with the fewest number of strokes (Costill et al., 1985). Performance was measured as the time spent in completing the 100-m freestyle event in an official competition. The 100-m freestyle was selected because it is the event in which most young swimmers participate on regular basis. It is also the most popular swimming event not only for young but also for adult/elite and master swimmers.

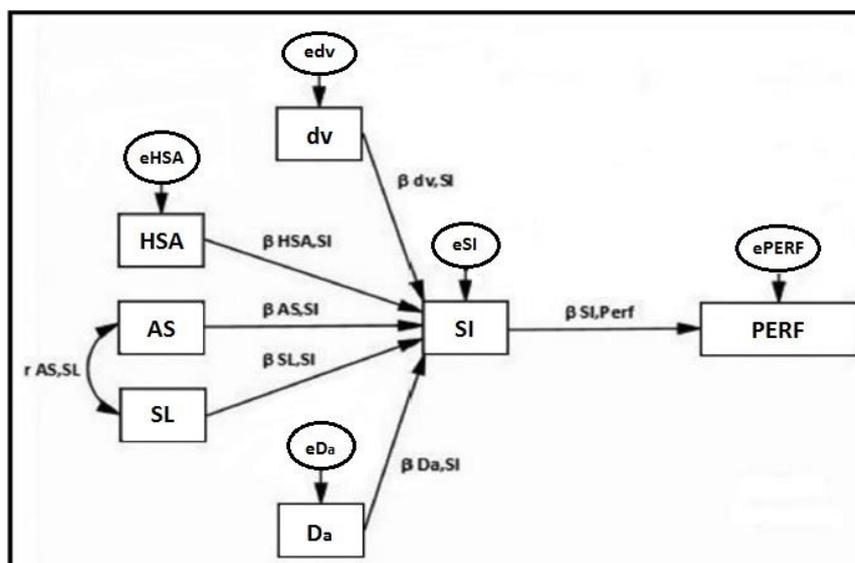


Figure 1. Theoretical path-flow model. AS – arm span; SL – stroke length;  $dv$  – intra-cyclic velocity fluctuation; HSA – hand surface area;  $D_a$  – active drag; SI – stroke index; PERF – performance;  $\beta_{x_i,y_i}$  – beta value for regression model between exogenous ( $x_i$ ) and endogenous ( $y_i$ ) variables;  $e_{x_i}$  – disturbance term for a given endogenous variable;  $r_{x_i,y_i}$  – correlation coefficient between two variables;  $x_i \rightarrow y_i$  – variable  $y_i$  depends from variable(s)  $x_i$ ;  $x_i \leftrightarrow y_i$  – variable  $y_i$  is associated to variable  $x_i$ .

### **Performance data collection**

Swimming performance was assessed against time lists of the 100-m freestyle event in short course competitions (i.e. 25-m swimming pool) at local, regional or national level competitions. The time gap between assessment of all variables and swimming performance was less than the two weeks reported in other studies on the relationships between swimming performance and kinematic and/or energetic variables in young swimmers (Barbosa et al., 2010a, Marinho et al., 2010a).

### **Anthropometric data collection**

The anthropometric variables selected for the path-flow model were the AS and the HSA. For the AS assessment, subjects were placed in an orthostatic position, with both arms in lateral abduction at a 90° angle with the trunk. Both arms and fingers were fully extended. The distance between the tip of each third finger was measured with a flexible anthropometric tape (RossCraft, Canada). The test/retest evaluation (i.e. Intraclass Correlation Coefficient) was very high for the AS (ICC = 0.99). For the HSA measurement, swimmers placed their dominant hand on the scan surface of a copy machine with fingers in the position they usually adopt while swimming. The scan surface was also fitted with a 2D calibration frame. Thereafter, the perimeter of the HSA was digitized in the Xerox machine (Xerox 4110, Norwalk, Connecticut, USA) and files were converted into pdf format. The HSA was afterward computed with dedicated software (Universal Desktop Ruler, v3.3.3268, AVPSOft, USA). The measurement procedures were: (i) scale calibration; (ii) digitization of hand surface perimeter and; (iii) computation and record of the HSA value (Morais et al., 2011). The test/retest evaluation was very high for the HSA (ICC = 0.99).

### **Biomechanical data collection**

Intra-cyclic velocity fluctuation, SL and SI were selected as kinematic variables. Each swimmer performed three bouts of 25-m freestyle from an underwater start. For further analysis the mean value of the three repetitions was computed. Subjects performed the bouts alone without other swimmers in the same swim lane or in nearby lanes to reduce drafting, pacing effects and bias in the drag force (Marinho et al., 2010a). The subjects were advised to reduce gliding after the start (Barbosa et al., 2010b). To assess  $dv$  a speedo-meter cable (Swim speedo-meter, Swimspotec,

Hildesheim, Germany) was attached to the swimmers' hip and the biosignal reading was acquired on-line at a sampling rate of 50 Hz. LabVIEW (v. 2009) software interface was used to acquire, display and process pairwise velocity-time data on-line during the swim bout. To transfer data from the speedo-meter to the software application a 12-bit resolution acquisition card (USB-6008, National Instruments, Austin, Texas, USA) was used (Barbosa et al., 2011). Data were exported to signal processing software (AcqKnowledge v. 3.5, Biopac Systems, Santa Barbara, USA) and filtered with a 5 Hz cut-off low-pass 4<sup>th</sup> order Butterworth filter. Intra-cyclic velocity fluctuation was computed as (Barbosa et al., 2010c):

$$dv = \frac{\sqrt{\sum_i (v_i - \bar{v})^2 F_i / n}}{\sum_i v_i F_i / n} \quad (1)$$

Where  $dv$  represents intra-cyclic velocity fluctuation (dimensionless),  $v$  represents the mean swimming velocity (in  $\text{m}\cdot\text{s}^{-1}$ ),  $v_i$  represents the instant swimming velocity (in  $\text{m}\cdot\text{s}^{-1}$ ),  $F_i$  represents the absolute frequency and  $n$  represents the number of observations. Stroke length was computed as (Craig and Pendergast, 1979):

$$SL = \frac{v}{SF} \quad (2)$$

Where  $SL$  represents stroke length (in m),  $v$  represents the mean swimming velocity (in  $\text{m}\cdot\text{s}^{-1}$ ) and  $SF$  represents the stroke frequency (in Hz). The  $v$  was calculated dividing the 13-m distance swam in the middle of the swimming pool by the time spent with a manual chronometer (Golfinho Sports MC 815, Aveiro, Portugal) by two expert evaluators (ICC = 0.97). The  $SF$  was measured with a chrono-frequency counter during three consecutive strokes by two expert evaluators (ICC = 0.96). Stroke index was also computed as a swim efficiency estimator (Costill et al., 1985):

$$SI = SL \cdot v \quad (3)$$

Where  $SI$  represents stroke index (in  $\text{m}^2\cdot\text{s}^{-1}$ ),  $SL$  represents stroke length (in m) and  $v$  is the mean swimming velocity (in  $\text{m}\cdot\text{s}^{-1}$ ).

### Hydrodynamic data collection

In the hydrodynamic domain, the  $D_a$  was computed using the velocity perturbation method (Kolmogorov and Duplisheva, 1992). Each swimmer performed two maximal 25-m bouts of freestyle with an underwater start. The first bout was performed

without the perturbation device and the second one with the perturbation device. Subjects performed the bouts alone without other swimmers in the same or nearby swim lanes to reduce drafting, pacing effects and bias in the drag force (Marinho et al., 2010a). Active drag was calculated from the difference between the swimming velocities both towing and without towing a perturbation buoy (additional hydrodynamic body; Kolmogorov and Duplisheva, 1992; Kolmogorov et al., 1997). The drag of the perturbation buoy was computed from the manufacturer's calibration of the buoy-drag characteristics and its velocity (Kolmogorov and Duplisheva, 1992). Swimming velocity was assessed over 13-m (between 11<sup>th</sup>-m and 24<sup>th</sup>-m) from the starting wall). The time spent to cover this distance was measured with a manual chronometer (Golfinho Sports MC 815, Aveiro, Portugal) by two expert evaluators as is customary with this method (Marinho et al., 2010a). The ICC for both evaluators was very high (ICC = 0.97). Active drag was calculated as (Kolmogorov and Duplisheva, 1992):

$$D_a = \frac{D_b v_b v^2}{v^3 - v_b^3} \quad (4)$$

Where  $D_a$  represents the swimmers' active drag at maximal velocity (in N),  $D_b$  is the resistance of the perturbation buoy (in N) and,  $v_b$  and  $v$  are the swimming velocities with and without the perturbation device (in  $m \cdot s^{-1}$ ), respectively.

### Statistical analysis

The Kolmogorov-Smirnov and the Levene tests were used to analyze normality and homocedasticity assumptions, respectively. Descriptive statistics (mean, one standard deviation, minimum and maximum) were computed. To assess the association between performance and remaining variables, Pearson correlation coefficients were computed between swimming performance and all selected variables ( $P \leq 0.05$ ). As rule of thumb, for qualitative and effect size assessments, the relationship was defined as: (i) very weak if  $r^2 < 0.04$ ; weak if  $0.04 \leq r^2 < 0.16$ ; moderate if  $0.16 \leq r^2 < 0.49$ ; high if  $0.49 \leq r^2 < 0.81$  and; very high if  $0.81 \leq r^2 < 1.0$ . The level of statistical significance was set at  $P \leq 0.05$ . For the structural equation modeling the path-flow analysis procedure was used. The interpretation of this kind of approach is based on: (i) the variables included (variables are inserted inside squares); (ii) the paths (i.e. arrows; an arrow between two variables means that one

variable determines the other); (iii) beta values (i.e. these suggest the contribution of one variable to the other: when the origin variable increases by one unit the destination variable increases by the amount of the beta value) and; (iv) residual errors and/or determination coefficient (represents the variable predictive error or the variable predictive value, respectively). Thereafter the model was computed and a confirmatory model obtained (i.e. a model that verified and confirmed the theoretical one). The estimation of linear regression standardized coefficients between exogenous and endogenous variables was computed. Standardized regression coefficients (b) were considered, and the significance of each one was assessed with the Student's *t* test ( $P \leq 0.05$ ). When a given path was significant ( $P \leq 0.05$ ) and with a moderate/strong association it was reported as being "meaningful" (Winter, 2008).

The quality of the model goodness-of-fit was measured by computing: (i) the ratio Chi-square/degrees of freedom ( $\chi^2/df$ ) and; (ii) the comparative fit index (CFI). The ratio Chi-square/degrees of freedom was considered qualitatively if (Wheaton, 1987):  $\chi^2/df > 5$  bad adjustment;  $5 \geq \chi^2/df > 2$  low adjustment;  $2 \geq \chi^2/df > 1$  good adjustment;  $\chi^2/df \sim 1$  very good adjustment. The comparative fit index was considered qualitatively if (Bentler, 1990):  $CFI < 0.90$  bad adjustment;  $0.90 \leq CFI < 0.95$  good adjustment;  $CFI \geq 0.95$  very good adjustment.

## Results

Table 1 presents descriptive statistics for overall sample (boys plus girls), boys only and girls only for all selected variables. Data variability, assessed by one standard deviation value, were moderate-high. This is especially obvious, concerning the overall statistics, for the HSA, ranging between 83.26 cm<sup>2</sup> and 163.84 cm<sup>2</sup>, for the *D<sub>a</sub>*, ranging between 11.81 N and 73.15 N, as well as for the swimming performance, ranging between 65.21 s and 128.30 s. For boys, the HSA ranged between 100.35 cm<sup>2</sup> and 163.84 cm<sup>2</sup>, the *D<sub>a</sub>* between 11.81 N and 73.15 N and swimming performance between 65.21 s and 106.18 s. For girls, the HSA, *D<sub>a</sub>* and swimming performance ranged between 83.26 cm<sup>2</sup> and 133.76 cm<sup>2</sup>, 16.49 N and 54.59 N, and 69.90 s and 128.30 s, respectively.

Table 2 presents the Pearson's correlation coefficients between swimming performance and remaining selected variables for overall total (boys plus girls), boys only and girls only. Data revealed that swimming performance was meaningfully

associated with SI (overall:  $r = -0.80$ ,  $P < 0.01$ ; boys:  $r = -0.87$ ,  $P < 0.01$ ; girls:  $r = -0.82$ ,  $P < 0.01$ ) and SL (overall:  $r = -0.64$ ,  $P < 0.01$ ; boys:  $r = -0.61$ ,  $P = 0.04$ ; girls:  $r = -0.61$ ,  $P = 0.02$ ). On the other hand, swimming performance was not significantly associated with the dv (overall:  $r = 0.18$ ,  $P = 0.39$ ; boys:  $r = -0.05$ ,  $P = 0.86$ ; girls:  $r = 0.13$ ,  $P = 0.64$ ) nor with the HSA (overall:  $r = -0.27$ ,  $P = 0.11$ ; boys:  $r = -0.17$ ,  $P = 0.57$ ; girls:  $r = -0.09$ ,  $P = 0.66$ ).

Table 1. Overall, boys and girls descriptive statistics of anthropometric, kinematic, hydrodynamic and swimming performance variables.

	Mean			1 SD			Minimum			Maximum		
	Overall	Boys	Girls	Overall	Boys	Girls	Overall	Boys	Girls	Overall	Boys	Girls
<b>AS [cm]</b>	156.36	161.56	152.47	11.72	12.99	9.26	135.00	141.00	135.00	187.00	187.00	170.00
<b>HSA [cm<sup>2</sup>]</b>	120.16	129.52	112.23	18.50	20.16	12.51	83.26	100.35	83.26	163.84	163.84	133.76
<b>dv [dimensionless]</b>	0.09	0.08	0.10	0.03	0.02	0.03	0.05	0.06	0.05	0.19	0.16	0.19
<b>SL [m]</b>	1.54	1.58	1.47	0.23	0.23	0.23	1.04	1.04	1.04	1.98	1.98	1.92
<b>D<sub>a</sub> [N]</b>	38.96	43.82	34.20	17.16	18.14	12.29	11.81	11.81	16.49	73.15	73.15	54.59
<b>SI [m<sup>2</sup>·c<sup>-1</sup>·s<sup>-1</sup>]</b>	1.92	2.06	1.79	0.47	0.56	0.36	0.90	0.90	1.01	2.86	2.86	2.45
<b>PERF [s]</b>	82.07	78.33	85.25	12.96	12.07	13.89	65.21	65.21	69.90	128.30	106.18	128.30

AS – arm span; HAS – hand surface area; dv – intra-cyclic velocity fluctuation; SL – stroke length; D<sub>a</sub> – active drag; SI – stroke index; PERF – performance.

Table 2. Overall, boys and girls Person's correlation coefficients between swimming performance and remain variables selected.

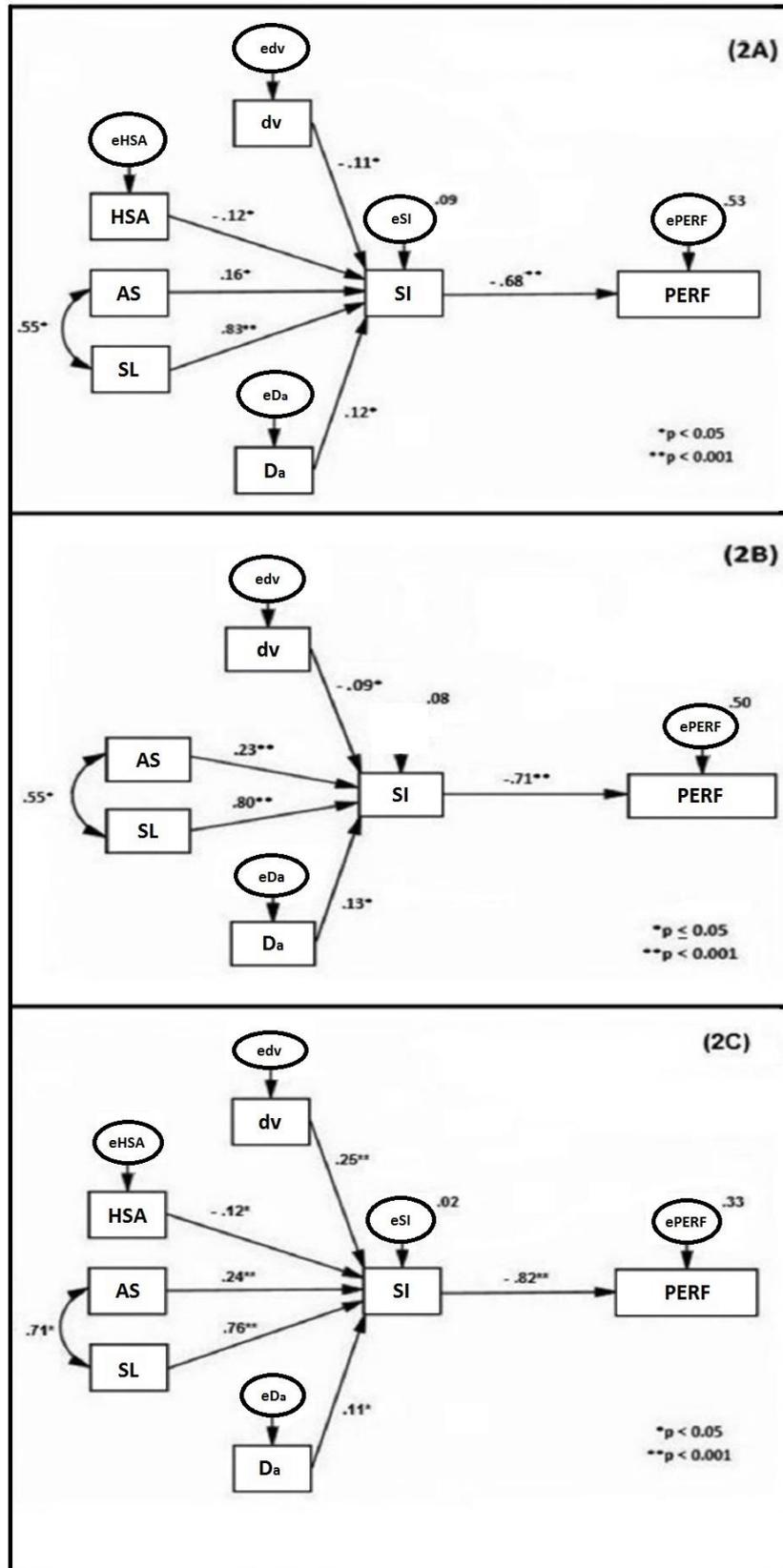
	r			P value		
	Overall	Boys	Girls	Overall	Boys	Girls
AS [cm]	- 0.35	- 0.73	- 0.16	0.03	0.006	0.479
HSA [cm <sup>2</sup> ]	- 0.27	- 0.17	- 0.10	0.11	0.575	0.668
dv [dimensionless]	0.18	- 0.06	0.14	0.39	0.865	0.642
SL [m]	- 0.64	- 0.61	- 0.61	0.001	0.045	0.02
D <sub>a</sub> [N]	- 0.45	- 0.62	- 0.49	0.03	0.137	0.066
SI [m <sup>2</sup> ·c <sup>-1</sup> ·s <sup>-1</sup> ]	- 0.80	- 0.88	- 0.83	0.001	0.001	0.001

AS – arm span; HSA – hand surface area; dv – intra-cyclic velocity fluctuation; SL – stroke length; D<sub>a</sub> – active drag; SI – stroke index.

Figure 2 presents the confirmatory path-flow models for young swimmers' performance (overall: 2A and 2B; boys: 2C and 2D; girls: 2E and 2F) based on selected anthropometric, hydrodynamic and kinematic variables. In each path the b value is reported (i.e., the standardized regression weight) for the regression model between each exogenous and endogenous variable. When the exogenous variable changes (i.e., origin of the path) by one unit, the endogenous variable (i.e. destination of the path) changes by the same quantity as the beta value. All paths linked in the theoretical model were significant in the confirmatory model. The overall model goodness-of-fit when including all variables (Figure 2A) was: (i)  $\chi^2/df = 7.058$  (i.e. bad adjustment) and; (ii) CFI = 0.601 (i.e. bad adjustment). For boys (Figure 2C) was: (i)  $\chi^2/df = 4.607$  (i.e. low adjustment) and; (ii) CFI = 0.592 (i.e. bad adjustment). For girls (Figure 2E) was: (i)  $\chi^2/df = 3.516$  (i.e. low adjustment) and; (ii) CFI = 0.640 (i.e. bad adjustment).

Deleting the HSA-SI path in the overall model (Figure 2B), for boys (Figure 2D) and for girls (Figure 2F) with subsequent recomputation of the remaining data, the new confirmatory model increased the predictive value of the models. The prediction of swimming performance based solely on biomechanics and its determining domains was 50%, 58% and 62% for overall data, boys and girls respectively. The SI was predicted based on remaining kinematic, anthropometric and hydrodynamic variables at 92%, 97% and 94% for overall data, boys and girls respectively. Moreover, the model goodness-of-fit improved meaningfully: (i)  $\chi^2/df = 1.908$  (i.e. good adjustment); CFI = 0.940 (i.e. good adjustment) for overall data; (ii)  $\chi^2/df = 1.612$  (i.e. good adjustment) and; CFI = 0.931 (i.e. good adjustment) for boys; (iii)  $\chi^2/df = 3.010$  (i.e.

low adjustment); CFI = 0.779 (i.e. bad adjustment) for girls. In this sense, the overall and boys confirmatory models had a good adjustment, although it was low for girls.



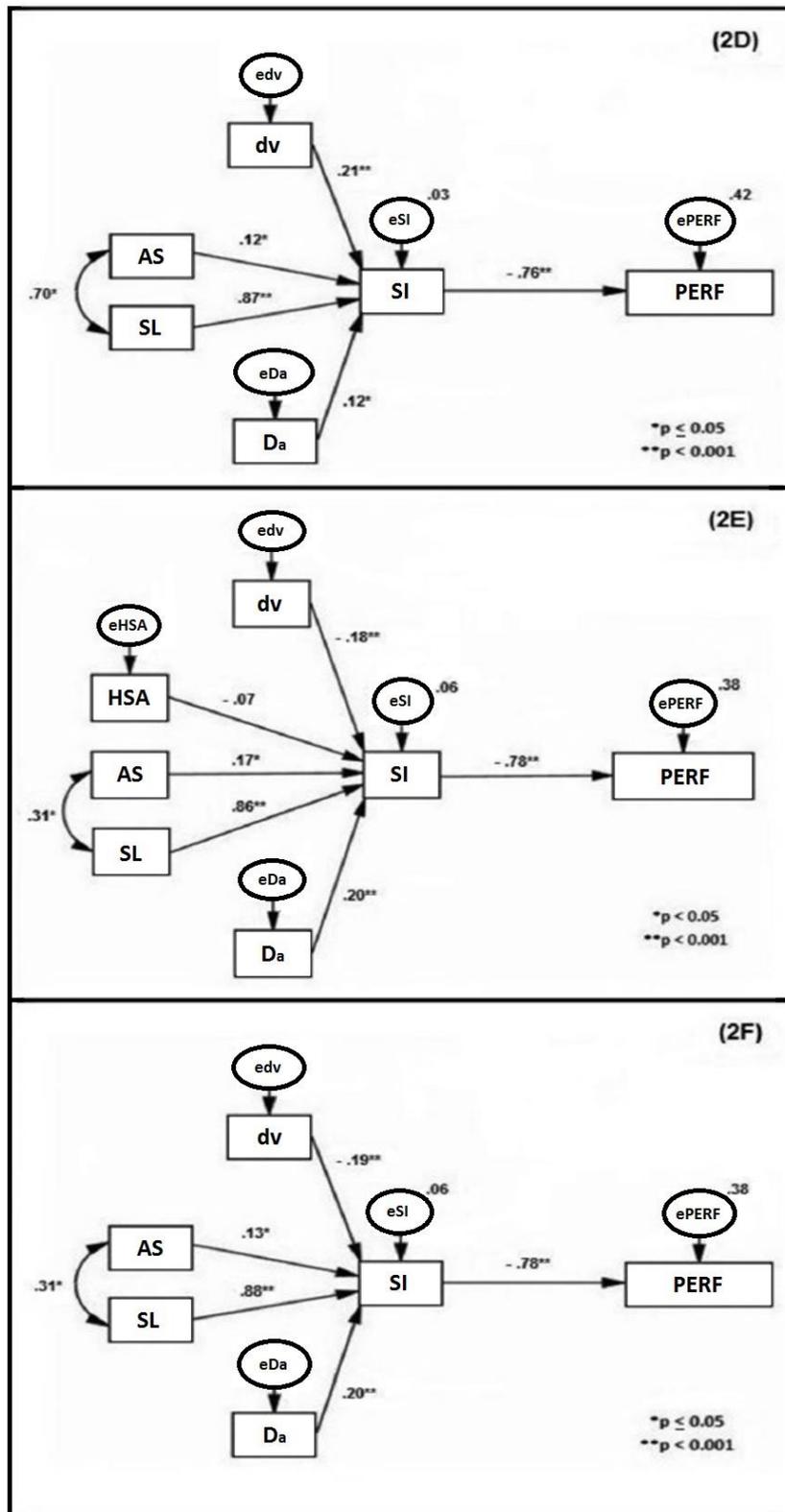


Figure 2. Overall confirmatory path-flow model including all variables computed (2A) and deleting variable that allowed to reduce the residual error and improve the goodness-of-fit (2B) with the subsequent recomputation of remain data. Boy's confirmatory path-flow model including all variables computed (2C) and deleting

variable that allowed to improve significantly the goodness-of-fit (2D). Girl's confirmatory path-flow model including all variables computed (2E) and deleting variable that allowed to improve the goodness-of-fit. HSA – hand surface area; AS – arm span; SL – stroke length; dv – intra-cyclic velocity fluctuation;  $D_a$  – active drag; SI – stroke index;  $x_i \rightarrow y_i$  – variable  $y_i$  depends from variable(s)  $x_i$ ;  $x_i \leftrightarrow y_i$  – variable  $y_i$  is associated to variable  $x_i$ .

## Discussion

The aim of this study was to develop a structural equation model for young swimmers' performance based on selected kinematic, anthropometric and hydrodynamic variables and to quantify the partial contribution of biomechanics to young swimmers' performance. Main data showed that swimming performance is dependent on SI (an efficiency estimator) and this in turn on the dv, SL, AS and  $D_a$ . The prediction of swimming performance based solely on biomechanics was very high ( $0.50 \leq r^2 \leq 0.62$ ).

Mean data reported are similar to other studies involving prepubescent swimmers (Barbosa et al., 2010a; Barbosa et al., 2010b; Latt et al., 2009a; Jurimae et al., 2007). To the best of our knowledge, the HSA has never been assessed in young swimmers, except for hand length and width (Vitor and Bohme, 2010), and hand size (Helmuth, 1980). The data revealed a moderate-high dispersion, namely for performance and the HSA, which allowed the analysis of hypothetical relationships between these selected variables and swimming performance over a broader scope. Pearson's correlation coefficients showed that swimming performance was significantly correlated with all variables, except for the dv and the HSA. The highest correlation values were for the SI and the SL. At least in adult/elite swimmers, higher-skilled swimmers present a higher SL than lower-skilled counterparts (Barbosa et al., 2010c). The SI is also higher in international level than in national level swimmers (Sánchez and Arellano, 2002). Scientific evidence for young swimmers is not so obvious, mainly because research with this cohort is scarce. However, it seems that the data for young swimmers is similar to that for their older counterparts. A higher AS is also associated with a higher performance level in young swimmers (Jurimae et al., 2007). Arm span imposes an increase in the SL (Pelayo et al., 1997). The  $D_a$  was also correlated with performance. To enhance performance, swimmers have to

increase swimming velocity, which is one of the main determinants of  $D_a$ . It was hypothesized that a higher HSA might increase propulsion. However, the correlation was not significant. Although an increased HSA might be an advantage, it should be stressed that the appropriate hand orientation (i.e. attack and pitch angles) on stroking has a role in enhancing performance (Rouboa et al., 2006; Bixler and Riewald, 2002). Probably some of the subjects assessed did not perform an appropriate hand orientation as well as varying in HSA.

The first overall confirmatory model (Figure 2A overall; boys 2C and girls 2E), including the HSA linked to SI, had a bad adjustment. Some studies suggested a relationship between hand shape (i.e. hand length) and swimming efficiency, or at any rate its thrust (Marinho et al., 2010b; Alves et al., 1998). However, most of those studies assessed adult swimmers (Gourgoulis et al., 2010) or made numerical simulations from adult models (Rouboa et al., 2006; Bixler and Riewald, 2002). There are few studies regarding its relationship in young swimmers, including pubescent ones (e.g. Alves et al., 1998; Helmuth, 1980). However, one study (Helmuth, 1980) reported a positive correlation between hand size and swimming performance in young swimmers. Despite this, it can be stated that there is no solid scientific evidence that at such early ages the HSA is as determinant of swimming performance or of swimming efficiency as it is in adult/elite swimmers.

The second confirmatory model (Figure 2B overall; boys 2D and girls 2F) removed the HSA-SI path presenting two hierarchical levels, and increased the model goodness-of-fit (i.e. good adjustment). The second level is the relationship between the SI and remaining kinematic, anthropometric and hydrodynamic variables selected. The SI is considered a viable variable by which to estimate overall swimming efficiency (Costill et al., 1985). The capacity to cover a given distance (i.e. SL) at greater velocity represents an increased swimming efficiency. The variables maintained in the final overall confirmatory model (i.e. AS, SL,  $dv$  and  $D_a$ ) had high ability to predict SI (overall:  $r^2 = 0.92$ ; boys:  $r^2 = 0.97$ ; girls:  $r^2 = 0.94$ ). From those variables, the SL had the higher standardized direct effect to SI (overall:  $\beta = 0.80$ ,  $P < 0.001$ ; boys:  $\beta = 0.87$ ,  $P < 0.001$ ; girls:  $\beta = 0.88$ ,  $P < 0.001$ ). This signifies that when SL increased by one meter, SI increased by  $0.80 \text{ m}^2 \cdot \text{s}^{-1}$ ,  $0.87 \text{ m}^2 \cdot \text{s}^{-1}$  and  $0.88 \text{ m}^2 \cdot \text{s}^{-1}$  overall, boys and girls respectively. This is obvious since the SI is computed on the

basis of SL and the swimming velocity. Arm span is usually reported as being related to swimming performance (Latt et al., 2009a) because it is associated with improved swim efficiency (Saavedra et al., 2010). Another viable method of analyzing the overall swim mechanics is by means of the swimmers'  $dv$ . Swimmers do not maintain a constant swim velocity due to variations of the limbs and trunk within the stroke cycle (Barbosa et al., 2008). Such a fact might decrease energy cost and thus improve swim efficiency. In this particular case, when  $dv$  increased by one arbitrary unit (a.u.), SI decreased by  $0.09 \text{ m}^2\cdot\text{s}^{-1}$  and  $0.19 \text{ m}^2\cdot\text{s}^{-1}$  for overall and girls models, respectively, though for boys it had essentially no effect. Active drag was also included in the model, since to maintain displacement, swimmers must overcome drag forces (Kjendlie and Stallman, 2008). To do this, they have to adopt the best possible hydrodynamic positions and segmental kinematics throughout the stroke.

The final confirmatory first level included the SI-performance relationship. The SI had a moderate-high standardized direct effect on performance (overall:  $\beta = -0.71$ ,  $P < 0.001$ ; boys:  $\beta = -0.76$ ,  $P < 0.001$ ; girls:  $\beta = -0.78$ ,  $P < 0.001$ ). Without considering other scientific domains, the biomechanical domain and its determinants were good predictors of the performance (overall:  $r^2 = 0.50$ ; boys:  $r^2 = 0.58$ ; girls:  $r^2 = 0.62$ ). A previous study (Barbosa et al., 2010a) predicted performance in roughly 80% of cases, based only on biomechanical and energetic domains. It was not the aim of this paper to replicate this study (Barbosa et al., 2010a), using the same variables. Instead, the goal was to expand the biomechanical "branch" of the model reported by (Barbosa et al., 2010a) and to identify the anthropometric and hydrodynamic determinants and to understand the interplay between them. Thus, it can be speculated that remaining 30% (to increase the performance prediction up to 80% as previously reported) might be attributable to energetics, a domain not considered here. It could therefore be interesting in future to develop the energetics "branch" of the original model. Indeed, most of the technical and scientific evidence for young swimmers suggests that the best way to enhance performance is through improving technique. Swimming efficiency should be the focus at these ages, more so than the energy profile or other fitness components such as muscle strength or anaerobic fitness (Garrido et al., 2010; Van Praagh, 2000). Our data also suggests that, for young swimmers, biomechanics may well have a higher performance prediction power than energetics. Therefore, technique should represent the core of the training

program at these ages. Coaches should therefore design training programs focusing on improvement swimming technique (i.e. increasing the swimming efficiency). In prior exploratory researches the SI was one of the best performance predictors (Vitor and Bohme, 2010; Latt et al., 2009a; Klika and Thorland, 1994). For these studies, the SI-performance ranged from moderate to very high associations.

Young swimmers' coaches and practitioners should thus design training programs with a focus on specific training sets for technique correction using a large variety of drills. By increasing swimming efficiency it is possible to meaningfully enhance the performance for this age-group. However, to increase swimming efficiency some further variables should be manipulated. Coaches must pay extra attention to technical issues such as an increased SL (related to a higher AS) and a better hydrodynamic position so as to decrease  $D_a$ . Emphasis should also be given to improving stroke mechanics (e.g. inter-limb coordination in opposition and/or superposition) to avoid swim discontinuities as observed in adult/elite swimmers (Seifert et al., 2010). This same logic ought also be applied in talent identification and selection programs. The main limitations of this research were as follows: (i) a direct measure of the propulsive efficiency was not adopted, merely a swim efficiency estimator; (ii) short distance events such as the 100-m freestyle are strongly associated with energetics variables, at least in adult/elite swimmers, but not so obviously in younger counterparts; (iii) not included in the model were variables related to functional fitness (e.g. muscular strength or flexibility) that might influence stroke mechanics.

## Conclusions

To conclude, it was possible to develop a confirmatory model to explain swimming performance in young swimmers. The data suggested that the biomechanical domain contributed 50% to overall sample performance (boys plus girls), 58% to boys only performance and 62% only to girls-only performance. Increasing swimming efficiency (i.e. improving swim technique) leads to a performance enhancement. On the other hand, swimming efficiency improvement is related to a decrease in the  $dv$  and an increase of the SL and AS. However, the increase in the  $D_a$  is a result of the increase in swimming velocity. It would appear that the best way to improve performance is to improve technique, thus increasing efficiency and optimizing hydrodynamic position.

Therefore, the focus of training sessions for young swimmers should be on the enhancement of technique.

## References

Alves F, Costa M, Gomes-Pereira J. (1998). The influence of swimming velocity on the kinematic characteristics of backstroke swimming. In: H Riehle, M Vieten (Eds), Proceedings of XVI International Symposium on Biomechanics in Sports (pp. 104-107). Konstanz: UVK Universitätsverlag Konstanz.

Barbosa TM, Fernandes RJ, Morouço P, Vilas-Boas JP. (2008). Predicting the intracyclic variation of the velocity of the centre of mass from segmental velocities in Butterfly stroke: A pilot study. *J Sports Sci Med.* 7: 201–209.

Barbosa TM, Costa M, Marinho DA, Coelho J, Moreira M, Silva AJ. (2010a). Modeling the links between young swimmer's performance: energetic and biomechanical profiles. *Pediatr Exerc Sci.* 22: 379–391.

Barbosa TM, Costa MJ, Marques MC, Silva AJ, Marinho DA. (2010b). A model for active drag force exogenous variables in young swimmers. *J Hum Sport Exerc.* 5(3): 379–388.

Barbosa TM, Bragada JA, Reis VM, Marinho DA, Carvalho C, Silva AJ. (2010c). Energetics and biomechanics as determining factors of swimming performance: Updating the state of the art. *J Sci Med Sport.* 13(2): 262–269.

Barbosa TM, Costa MJ, Morais JE, Jesus S, Silva AJ, Batista J, Gonçalves J. (2011). Validation with videometry of an integrated system to assess horizontal intra-cyclic velocity with a mechanical speedo-meter. In: JP Vilas-Boas, L Machado, W Wangdo, AP Veloso (Eds), *Biomechanics in Sports 29* (pp. 833–835). Porto: Portuguese Journal of Sport Sciences.

Benjanuvattra N, Blanksby BA, Elliott BC. (2001). Morphology and hydrodynamics resistance in young swimmers. *Pediatr Exerc Sci.* 13: 246–255.

Bentler PM. (1990). Comparative fit indexes in structural models. *Psychol Bull.* 107: 238–246.

Bixler BS, Riewald S. (2002). Analysis of swimmer's hand and arm in steady flow conditions using computational fluid dynamics. *J Biomech.* 35: 713–717.

Craig A, and Pendergast D. (1979). Relationships of stroke rate, distance per stroke and velocity in competitive swimming. *Med Sci Sports Exerc.* 11: 278–283.

Costill D, Kovalski J, Porter D, Fielding R, King D. (1985). Energy expenditure during front crawl swimming: predicting success in middle-distance events. *Int J Sports Med.* 6: 266–270.

Denadai B, Greco C, Teixeira M. (2000). Blood lactate response and critical speed in swimmers aged 10-12 years of different standards. *J Sports Sci.* 10: 779–784.

Duché P, Falgairette G, Bedu M, Lac G, Robert A, Coudert J. (1993). Analysis of performance of prepubertal swimmers assessed from anthropometric and bio-energetic characteristics. *Eur J Appl Physiol Occup Physiol.* 66(5): 467–471.

Garrido N, Marinho DA, Barbosa TM, Costa AM, Silva AJ, Pérez-Turpin JÁ, Marques MC. (2010). Relationships between dry land strength, power variables and short sprint performance in young competitive swimmers. *J Hum Sport Exerc.* 5(2): 240–249.

Geladas ND, Nassis GP, Pavlicevic S. (2005). Somatic and physical traits affecting sprint swimming performance in young swimmers. *Int J Sports Med.* 26: 139–144.

Gourgoulis V, Antoniou P, Aggeloussis N, Mavridis G, Kasimatis P, Vezos N, Boli A, Mavromatis G. (2010). Kinematic characteristics of the stroke and orientation of the hand during front crawl resisted swimming. *J Sports Sci.* 11: 1165–1173.

Greco C, Pelarigo JG, Figueira TR, Denadai BS. (2007). Effects of gender on stroke rates, critical speed and velocity of a 30-min swim in young swimmers. *J Sports Sci Med.* 6: 441–447.

Helmuth HS. (1980). Anthropometric survey of young swimmers. *Anthropol Anz.* 38(1): 17–34.

Hohmann A, and Seidel I. (2010). Talent prognosis in young swimmers. In: P Kjendlie, RK Stallman, J Cabri (Eds), *Biomechanics & Medicine in Swimming XI* (pp. 262-264). Oslo: Norwegian School of Sport Science.

Jurimae J, Haljaste K, Cicchella A, Latt E, Purge P, Leppik A, Jurimae T. (2007). Analysis of swimming performance from physical, physiological and biomechanical parameters in young swimmers. *Pediatr Exerc Sci.* 19: 70–81.

Kjendlie PL, and Stallman RK. (2008). Drag characteristics of competitive swimming children and adults. *J Appl Biomech.* 24(1): 35–42.

Klika RJ, and Thorland WG. (1994). Physiological determinants of sprint swimming performance in children and young adults. *Pediatr Exerc Sci.* 6: 59–68.

Kolmogorov S, and Duplishcheva O. (1992). Active drag, useful mechanical power output and hydrodynamic force in different swimming strokes at maximal velocity. *J Biomech.* 25:311–318.

Kolmogorov S, Rumyantseva O, Gordon B, Cappaert JM. (1997). Hydrodynamic characteristics of competitive swimmers of different genders and performance levels. *J Appl Biomech.* 13:88–97.

Latt E, Jürimäe J, Haljaste K, Cicchella A, Purge P, Jurimae T. (2009a). Longitudinal development of physical and performance parameters during biological maturation of young male swimmers. *Percept Mot Skills.* 108:297–307.

Latt E, Jurimae J, Haljaste K, Cicchella A, Purge P, Jurimae T. (2009b). Physical development and swimming performance during biological maturation in young female swimmers. *Coll Antropol.* 33: 117–122.

Latt E, Jurimae J, Maestu J, Purge P, Ramson R, Haljaste K, Keskinen K, Rodriguez FA, Jurimae T. (2010). Physiological, biomechanical and anthropometrical predictors of sprint swimming performance in adolescent swimmers. *J Sports Sci Med.* 9: 398–404.

Marinho DA, Barbosa TM, Costa MJ, Figueiredo C, Reis VM, Silva AJ, Marques MC. (2010a). Can 8 weeks of training affect active drag in young swimmers? *J Sports Sci Med.* 9: 71–78.

Marinho DA, Barbosa TM, Reis VM, Kjendlie PL, Alves FB, Vilas-Boas JP, Machado L, Ailva AJ, Rouboa A. (2010b). Swimming propulsion forces are enhanced by a small finger spread. *J Appl Biomech.* 26: 87–92.

Morais JE, Costa MJ, Mejias EJ, Marinho DA, Silva AJ, Barbosa TM. (2011). Morphometric study for estimation and validation of trunk transverse surface area to assess human drag force on water. *J Hum Kinet.* 28: 5–13.

Pelayo P, Wille F, Sidney M, Berthoin S, Lavoie JM. (1997). Swimming performances and stroking parameters in non skilled grammar school pupils: relation with age, gender and some anthropometric characteristics. *J Sports Med Phys Fitness.* 37(3): 187–193.

Poujade B, Hautier CA, Rouard A. (2002). Determinants of the energy cost of frontcrawl swimming in children. *Eur J Appl Physiol.* 87:1–6.

Rouboa A, Silva A, Leal L, Rocha J, Alves F. (2006). The effect of swimmer's hand/forearm acceleration on propulsive forces generation using computational fluid dynamics. *J Biomech.* 39(7): 1239–1248.

Saavedra JM, Escalante Y, Ferran AR. (2010). A multivariate analysis of performance in young swimmers. *Pediatr Exerc Sci.* 22: 135–151.

Sánchez J, and Arellano R. (2002). Stroke index values according to level, gender, swimming style and event race distance. In: K Gianikellis (Ed), *Proceedings of the XXth International Symposium on Biomechanics in Sports* (pp. 56-59). Cáceres: Universidad de Extremadura.

Seifert L, Toussaint HM, Alberty M, Schnitzler C, Chollet D. (2010). Arm coordination, power, and swim efficiency in national and regional front crawl swimmers. *Hum Mov Sci.* 29: 426–439.

Sidney M, Alberty M, Leblanc H, Chollet D. (2010). Stroking parameters during competition. In: L Seifert, D Chollet, I Mujika (Eds), *The world book of swimming: from science to performance* (pp. 443-458). New York: Nova Science Publishers.

Silva AJ, Costa AM, Oliveira PM, Reis VM, Saavedra J, Perl J, Rouboa A, Marinho DA. (2007). The use of neural network technology to model swimming performance. *J Sports Sci Med.* 6: 117–125.

Van Praagh E. (2000). Development of anaerobic function during childhood and adolescence. *Pediatr Exerc Sci.* 12: 150–173.

Vilas-Boas JP, Costa L, Fernandes RJ, Ribeiro J, Figueiredo P, Marinho DA, Silva AJ, Rouboa A, Machado L. (2010). Determination of the drag coefficient during the first and second gliding positions of the breaststroke underwater stroke. *J Appl Biomech.* 26: 324–331.

Vitor FM, and Böhme MT. (2010). Performance of young male swimmers in the 100 meters front crawl. *Pediatr Exerc Sci.* 22: 278–287.

Wheaton B. (1987). Assesement of fit in overidentified models with latent variables. *Sociol Methods Res.* 16: 118–154.

Winter E. (2008). Use and misuse of the term “significant”. *J Sports Sci.* 26(5): 429–430.

---

## **Chapter 2**

**Longitudinal modeling in sports: young swimmers' performance and biomechanics profile**

---



## Abstract

New theories about dynamical systems highlight the multifactorial interplay between determinant factors to achieve higher sports performances, including in swimming. Longitudinal research does provide useful information on the sportsmen's changes and how training help him to excel. These questions may be addressed in one single procedure such as latent growth modeling. The aim of the study was to model a latent growth curve of young swimmers' performance and biomechanics over a season. Fourteen boys ( $12.33 \pm 0.65$  years-old) and 16 girls ( $11.15 \pm 0.55$  years-old) were evaluated. Performance, stroke frequency, intra-cyclic velocity fluctuation, arm's propelling efficiency, active drag, active drag coefficient and power to overcome drag were collected in four different moments of the season. Latent growth curve modeling was computed to understand the longitudinal variation of performance (endogenous variables) over the season according to the biomechanics (exogenous variables). Latent growth curve modeling showed a high inter- and intra-subject variability in the performance growth. Gender had a significant effect at the baseline and during the performance growth. In each evaluation moment, different variables had a meaningful effect on performance (M1:  $D_a$ ,  $\beta = -0.62$ ; M2:  $D_a$ ,  $\beta = -0.53$ ; M3:  $\eta_p$ ,  $\beta = 0.59$ ; M4: SF,  $\beta = -0.57$ ; all  $P < 0.001$ ). The models' goodness-of-fit was  $1.40 \leq \chi^2/df \leq 3.74$  (good-reasonable). Latent modeling is a comprehensive way to gather insight about young swimmers' performance over time. Different variables were the main responsible for the performance improvement. A gender gap, intra- and inter-subject variability was verified.

**Key-words:** *modeling, kinematics, hydrodynamics, season adaptations, contribution*



## Introduction

Talent identification, development, and follow-up are some of the major challenges that sports researchers and practitioners still face nowadays. Swimming performance is characterized by the multi-dimensional interplay of different scientific fields, where a highly complex interaction between several variables exists (Barbosa et al., 2010). Cross-sectional studies reported relationships between young swimmers' performance, Energetics (Toubekis et al., 2011), Biomechanics (Morais et al., 2012) and Motor Control (Silva et al., 2013). Nevertheless, from among all these scientific fields, Biomechanics plays a major role by explaining 50–60% of the performance of young swimmers (Morais et al., 2012). Probably the partial contribution of each key factor to performance may change across time, for example, over a season. However, until now no longitudinal research has been conducted about it in sports performance.

Moreover, longitudinal research should help in gathering insight into: (i) how biomechanical variables interplay and affect performance; (ii) the dynamical changes that happen at these early ages; (iii) the partial contribution of each determinant factor over time. For a long time sports research was based on the assumption that intra- and inter-subject variability should be minimized. Nowadays, dynamic systems theory and non-linear approaches suggest that variability should not be considered as a random error (Seifert et al., 2013). Evidence has been gathered lately about this topic in adult/elite swimmers (Costa et al., 2013; Komar et al., 2014) even though definitive answers are needed. Besides this, little or almost nothing is known about it in young swimmers. Interestingly young sportsmen, including swimmers, are supposed to be among the ones with a higher variability due to their allegedly low expertise level. It seems that athletes with lower (such as young swimmers) and very high expertise (including elite swimmers) levels are the ones with the highest variability (Seifert et al., 2011).

Until now, classical research designs and data analysis procedures (e.g. analysis of variance and regression models) selected on regular basis in sports performance were not helpful in gathering insight about such highly dynamic and complex relationships. Latent growth curve modeling is a structural equation modeling technique for longitudinal dataset. It is characterized by estimating intra- and inter-

subject growth trajectories, enabling researchers to predict future development (Wu et al., 2009). Structural equation modeling also allows the quantification of how much an exogenous variable contributes to an endogenous variable (Morais et al., 2012). Hence, its potential to explain complex and dynamic changes as reported earlier should be explored. This longitudinal data analysis procedure is reported on regular basis in Social Sciences such as Psychology (Castellanos-Ryan et al., 2013; Biesanz et al., 2003). In Sport Sciences a couple of papers can be found on physical fitness and health (Park & Schutz, 2005; Maia et al., 2003) but it was never attempted in sports performance as much as we are aware of.

Therefore, the aim of this study was to model a latent growth curve of young swimmers' performance and biomechanics over a season. It was hypothesized that latent growth curve modeling would explain performance improvement. Different exogenous variables would have a higher contribution on the performance enhancement throughout the season with a significant gender effect.

## **Methods**

### **Subjects**

Thirty young swimmers, including 14 boys:  $12.33 \pm 0.65$  years,  $284.85 \pm 67.48$  FINA (Fédération Internationale de Natation) points at the short-course meter (i.e. 25-m length swimming pool) 100-m freestyle; and 16 girls:  $11.15 \pm 0.55$  years,  $322.56 \pm 45.18$  FINA points at the short-course meter 100-m freestyle were recruited. All swimmers were in Tanner stages 1-2 by self-report at baseline (Tanner, 1962). The sample included age-group national record holders and champions. The swimmers were part of a national talent ID scheme. At the beginning of the research the swimmers had  $3.40 \pm 0.56$  years of training experience. Figure 1 reports the external training load over the season. Coaches, parents, and/or guardians consented and the athletes assented their participation on this study. All procedures were in accordance to the Helsinki Declaration regarding Human Research. The University of Trás-os-Montes and Alto Douro Ethic Committee also approved the study design (ethic review: UTAD-2011-219).

## Study design

The research design (Figure 2) included repeated measures of kinematic and hydrodynamic variables in four different moments over one season (i.e. longitudinal research). Testing sessions happened immediately before the beginning of the season (baseline-M1), 4 weeks later (first competition-M2), in the middle of the season (24<sup>th</sup> week-M3) and at the end of the season (38<sup>th</sup> week-M4). Data collection procedures were carried out in the same conditions at all times (e.g. the same swimming pool, lane, time of day).

## Theoretical model

Theoretical model (Figure 3) was designed to include kinematic, hydrodynamic, and performance, controlling the gender effect. Stroke frequency (SF), intra-cyclic velocity fluctuation ( $dv$ ) and propelling efficiency ( $\eta_p$ ) were selected as kinematic outcomes. As for hydrodynamics, active drag ( $D_a$ ), coefficient of active drag ( $C_{Da}$ ) and power to overcome drag ( $P_d$ ) were selected. Literature reports that kinematics and hydrodynamics determine young swimmers' performance (Morais et al., 2012). Stroke frequency, intra-cyclic velocity fluctuation, and arm's propelling efficiency (i.e. kinematics), active drag, coefficient of active drag, and power to overcome drag (i.e. hydrodynamics) are some of the variables that have a strong relationship with young swimmers' performance and therefore were selected on regular basis in swimming research (Silva et al., 2013; Morais et al., 2012; Marinho et al., 2010).

Swimming performance was chosen as the main outcome (endogenous variable; i.e. dependent variable being predicted), because the primary goal of coaches and swimmers is to enhance the performance. Kinematic and hydrodynamic variables are the exogenous variables (i.e. independent variables that predict the main outcome). The interpretation of this kind of approach is based on: (i) the variables included (inserted inside squares); (ii) the paths (i.e. arrows; an arrow between two variables means that one variable determines the other); (iii) beta values (i.e. these suggest the contribution of one variable to the other; when the origin variable increases by one unit the destination variable increases by the amount of the beta value); (iv) residual errors and/or determination coefficient (represents the variable predictive error or the variable predictive value, respectively, in the linked ellipse), and (v) the

latent variables (inserted in ellipses) are the no-observed (i.e. the slope analyzes the endogenous variable growth and variability; the intercept analyzes the variability in the baseline).

It was possible to extract the following details from the model: (i) the direct effect (i.e. contribution) of an exogenous variable to the endogenous one (i.e. performance) in each evaluation moment; (ii) the longitudinal growth of the endogenous variable; and (iii) the gender influence at the baseline values (intercept) and also in the endogenous variable growth (i.e. slope).

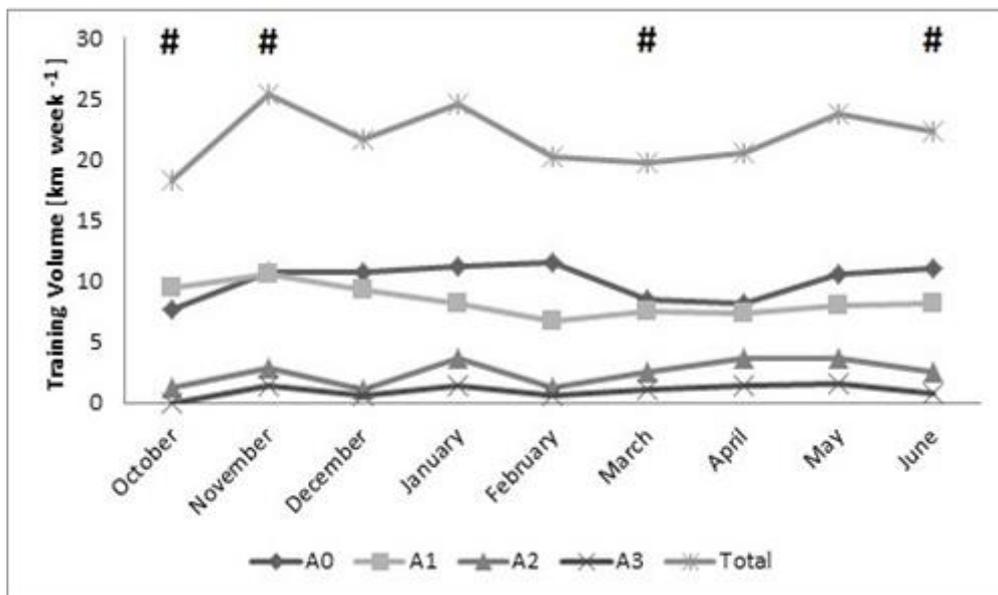


Figure 1. Training volume per week (in km) throughout the season. # – evaluation moments (Mi); A0 – warm-up and recovery pace; A1 – slow pace; A2 – moderate pace (aerobic capacity); A3 – intense pace (aerobic power).

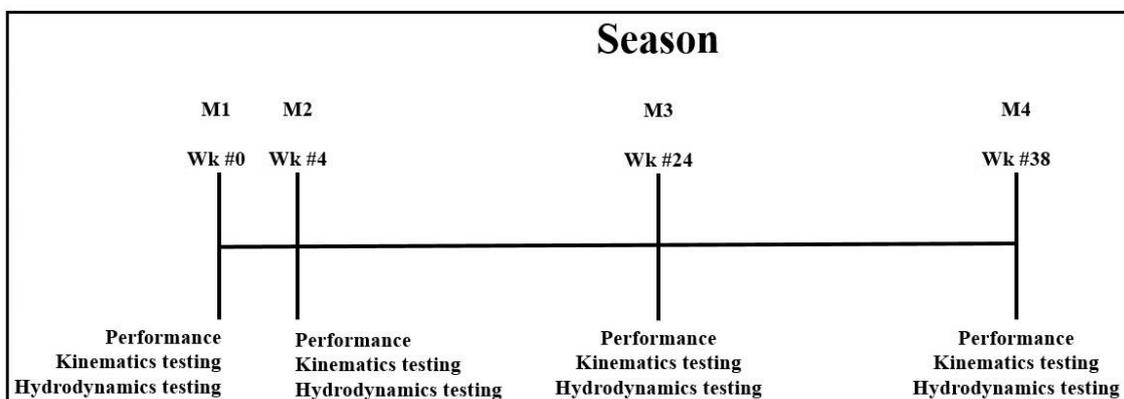


Figure 2. Study design scheme. M – moment; Wk – week; # – week's number.

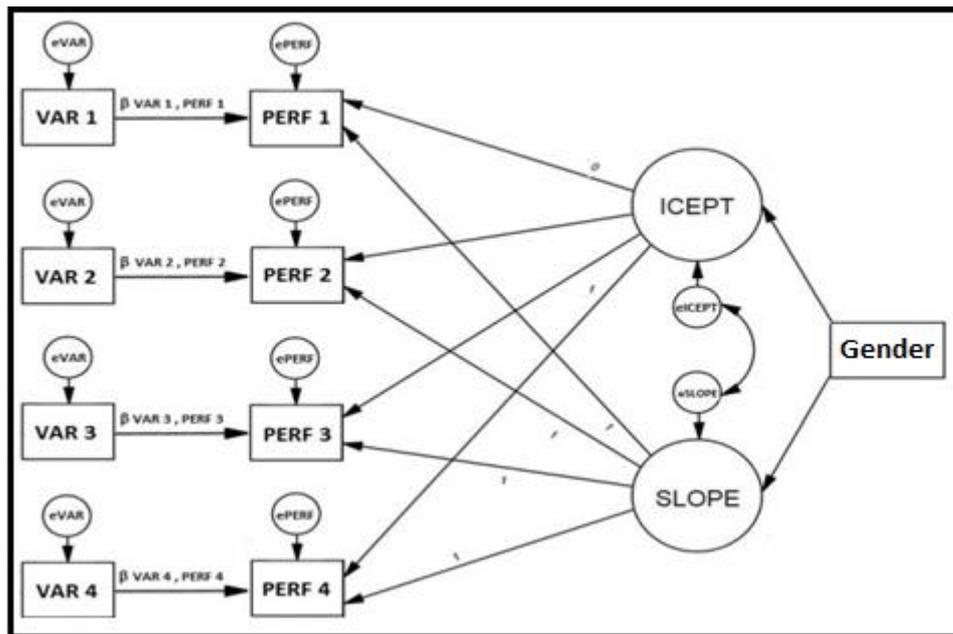


Figure 3. Theoretical model. VAR (1, 2, 3 and 4) – exogenous variable in M1, M2, M3 and M4, respectively; PERF (1, 2, 3 and 4) – performance in M1, M2, M3 and M4, respectively; ICEPT – intercept effect; SLOPE – slope effect; Gender – gender effect;  $\beta_{x_i, y_i}$  – beta value for regression model between exogenous ( $x_i$ ) and endogenous ( $y_i$ ) variables;  $e_{x_i}$  – disturbance term for a given variable;  $x_i \rightarrow y_i$  – variable  $y_i$  depends from variable  $x_i$ .

### Performance data collection

The official short course 100-m freestyle race was chosen as performance variable. The time gap between each the race and data collection took no longer than 15 days.

### Kinematics data collection

Swimmers were instructed to perform three maximal trials of 25-m at front-crawl with push-off start. Between each trial they had a 30-min rest to ensure full recovery. For further analysis the average value of the three trials was calculated (ICC = 0.96).

Kinematic data was collected with a mechanical technique (Swim speedo-meter, Swimsportec, Hildesheim, Germany). A 12-bit resolution acquisition card (USB-6008, National Instruments, Austin, Texas, USA) was used to transfer data ( $f = 50$  Hz) to a customized software (LabVIEW® interface, v.2009) (Barbosa et al., 2010). Data were exported to a signal processing software (AcqKnowledge v.3.9.0, Biopac Systems,

Santa Barbara, USA) and filtered with a 5 Hz cut-off low-pass 4<sup>th</sup> order Butterworth filter. Intra-cyclic velocity fluctuation was computed as (Barbosa et al., 2010):

$$dv = \frac{\sqrt{\sum_i (v_i - \bar{v})^2 F_i / n}}{\sum_i v_i F_i / n} \quad (1)$$

where  $dv$  is the intra-cyclic velocity fluctuation (dimensionless),  $v$  is the mean swimming velocity (in  $\text{m}\cdot\text{s}^{-1}$ ),  $v_i$  is the instant swimming velocity (in  $\text{m}\cdot\text{s}^{-1}$ ),  $F_i$  is the absolute frequency and  $n$  is the number of observations per stroke cycle. Two expert evaluators measured the SF with a stroke counter (base 3) and then converted to SI units (ICC = 0.98). The  $\eta_p$  was estimated as (Zamparo et al., 2005):

$$\eta_p = \left[ \left( \frac{v \cdot 0.9}{2\pi \cdot \text{SF} \cdot l} \right) \cdot \frac{2}{\pi} \right] \cdot 100 \quad (2)$$

where  $\eta_p$  is the arm's propelling efficiency,  $v$  is the swimming velocity, SF is the stroke frequency and  $l$  is the distance between shoulder and tip of the 3<sup>rd</sup> finger during the insweep.

### Hydrodynamics data collection

The Velocity Perturbation Method was selected to assess the hydrodynamic variables (Kolmogorov and Duplisheva, 1992). Swimmers performed two extra maximal trials of 25-m at front crawl with push-off start (one trial with and the other without carrying on the perturbation device). Swimming velocity was assessed between the 11<sup>th</sup>-m and 24<sup>th</sup>-m from the starting wall (Marinho et al., 2010). The time spent to cover this distance was measured with a manual stopwatch (Golfinho Sports MC 815, Aveiro, Portugal) by two expert evaluators (ICC = 0.97). The evaluators followed the swimmer to have a good line of sight when the swimmer passed the two distance marks. The  $D_a$  was estimated as (Kolmogorov and Duplisheva, 1992):

$$D_a = \frac{D_b v_b v^2}{v^3 - v_b^3} \quad (3)$$

where  $D_a$  is the swimmers' active drag at maximal velocity (in N),  $D_b$  is the resistance of the perturbation buoy (in N) provided by the manufacturer,  $v_b$  and  $v$  are the swimming velocities with and without the perturbation device (in  $\text{m}\cdot\text{s}^{-1}$ ). The  $C_{D_a}$  was calculated as (Kolmogorov and Duplisheva, 1992):

$$C_{Da} = \frac{2 \cdot D_a}{\rho \cdot S \cdot v^2} \quad (4)$$

where  $C_{Da}$  is the active drag coefficient (dimensionless),  $\rho$  is the water density (assumed to be  $1000 \text{ kg}\cdot\text{m}^{-3}$ ),  $v$  is the mean swimming velocity (in  $\text{m}\cdot\text{s}^{-1}$ ) and  $S$  is the swimmers' projected frontal surface area ( $\text{m}^2$ ). The  $P_d$  was obtained from (Kolmogorov and Duplisheva, 1992):

$$P_d = D \cdot v \quad (5)$$

where  $P_d$  is the power to overcome drag (in  $W$ ),  $D$  is the drag (in  $N$ ) and  $v$  is the mean swimming velocity (in  $\text{m}\cdot\text{s}^{-1}$ ).

### Statistical procedures

The normality and homocedasticity assumptions were analyzed with the Shapiro-Wilk and the Levene tests, respectively. Descriptive statistics included the calculation of the mean, median, minimum, maximum and one standard deviation.

Latent growth curve modeling was used to compute the longitudinal variation of the swimmers' performance over the season. This technique is characterized by estimating intra-individual (represented by the growth parameters; i.e. intercept and slope for growth) in the inter-individual (differences between subjects) growth trajectories (Wu et al., 2009). The intercept and slope are latent variables, which means that they are not directly observed but rather inferred. The intercept determines where the participants' baseline is and how they differ in that specific moment, showing the inter-individual differences between the participants at the baseline, corresponding to M1 in this model). The slope is the average rate of growth, related to the variation throughout a time-frame. It shows the hypothetical differences between the observed moments, and if an inter-individual variability exists or not.

The effect between exogenous ( $SF$ ,  $dv$ ,  $\eta_p$ ,  $D_a$ ,  $C_{Da}$ , and  $P_d$ ) and endogenous (performance) variables was also considered. Endogenous variable is the one being predicted and the growth rate analyzed. Exogenous variables are the ones with a direct effect on performance in each evaluation moment. Path-flow analysis model was used to estimate the linear regression standardized coefficients between

exogenous and endogenous variables. Standardized regression coefficients ( $\beta$ ) were selected, and the significance of each one assessed with Student's t test ( $P \leq 0.05$ ).

The models' goodness-of-fit were measured with the ratio Chi-square/degrees of freedom ( $\chi^2/df$ ) (Wheaton, 1987). As a rule of thumb if:  $5 < \chi^2/df$  the model has a poor adjustment;  $2 < \chi^2/df \leq 5$  reasonable adjustment;  $1 < \chi^2/df \leq 2$  good adjustment;  $\chi^2/df \sim 1$  very good adjustment.

## Results

Performance improved between the first (M1,  $72.05 \pm 5.33s$ ) and last (M4,  $66.13 \pm 5.16s$ ) evaluation moments.  $D_a$  and  $P_d$  showed the highest variations across the season (Table 1). Some selected variables ( $\eta_p$ ,  $D_a$ ,  $C_{D_a}$  and  $P_d$ ) increased in a non-linear fashion way.

In M2 and M3, performance achieved 59% ( $P < 0.001$ ) and 99% ( $P < 0.001$ ) of the last evaluation (M4) (Figure 4). The slope variance was significant for all models, suggesting a heterogeneous growth rate of the performance and hence an inter-subject variability for the pooled sample (i.e. boys plus girls). The dv model was the one presenting the highest slope ( $\beta = 6.56$ ;  $P = 0.003$ ). The intercept variance was significant for all models computed, suggesting an inter-subject variability at the baseline for the pooled sample and  $\eta_p$  showed the highest intercept ( $\beta = 28.15$ ;  $P < 0.001$ ). Overall it seems that each participant had its own and unique growth rate suggesting a high inter-subject variability.

Table 1. Descriptive statistics for selected kinematic and hydrodynamic variables in each evaluation moment.

	PERF [s]				SF [Hz]				dv [dimensionless]				$\eta_p$ [%]			
	M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4
<b>Mean</b>	72.05	68.91	66.44	66.13	0.85	0.84	0.86	0.86	0.09	0.09	0.08	0.08	30.21	30.38	33.13	31.03
<b>1SD</b>	5.33	5.43	5.33	5.16	0.09	0.08	0.10	0.08	0.02	0.03	0.02	0.01	2.94	2.94	4.55	5.13
<b>Median</b>	72.02	69.32	67.43	66.66	0.83	0.83	0.83	0.83	0.09	0.09	0.08	0.08	30.47	30.78	32.12	30.73
<b>Minimum</b>	60.30	57.03	58.01	57.36	0.67	0.70	0.72	0.73	0.06	0.06	0.06	0.05	25.16	24.24	26.00	24.12
<b>Maximum</b>	81.00	79.12	76.85	76.06	1.05	0.99	1.07	1.03	0.16	0.17	0.14	0.10	35.80	35.06	44.99	53.06

	$D_a$ [N]				$C_{Da}$ [dimensionless]				$P_d$ [W]			
	M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4
<b>Mean</b>	45.84	71.01	83.25	75.62	0.40	0.49	0.58	0.50	68.65	99.04	116.93	109.62
<b>1SD</b>	27.76	33.70	37.30	36.96	0.22	0.17	0.26	0.22	47.12	49.90	53.83	57.08
<b>Median</b>	34.29	58.36	70.85	77.05	0.34	0.48	0.51	0.46	49.22	81.72	97.36	107.25
<b>Minimum</b>	22.20	26.64	22.20	20.66	0.16	0.22	0.18	0.18	31.34	36.55	30.44	24.00
<b>Maximum</b>	139.53	167.33	172.93	189.52	1.41	0.89	1.28	1.32	241.54	229.86	244.27	295.41

$M_i$  – evaluation moment; PERF – performance; SF – stroke frequency; dv – intra-cyclic velocity fluctuation;  $\eta_p$  – arm's propelling efficiency;  $D_a$  – active drag;  $C_{Da}$  – active drag coefficient;  $P_d$  – power to overcome drag.

Gender had a significant effect on the performance growth with significant paths to intercept and slope for all models (Figure 4). Both  $\eta_p$  and  $C_{Da}$  models presented the highest significant paths ( $\beta = 0.94$ ;  $P < 0.001$ ). Data showed that boys presented better performances than girls. The  $P_d$  model had the highest significant path ( $\beta = 0.86$ ;  $P < 0.001$ ).

All selected variables presented a significant direct effect on performance at least in one evaluation moment (Figure 4). In M1 the  $D_a$  presented the highest direct effect on performance ( $\beta = -0.62$ ;  $P < 0.001$ ; by each 1N increase, performance improved 0.62s). In M2 was once again the  $D_a$  ( $\beta = -0.53$ ;  $P < 0.001$ ), in M3 the  $\eta_p$  ( $\beta = 0.59$ ;  $P < 0.001$ ) and in M4 the SF ( $\beta = -0.57$ ;  $P < 0.001$ ). Hence, swimmers relied on different exogenous variables to enhance performance in different moments of the season.

The models' goodness-of-fit ranged between  $1.40 \leq x^2/df \leq 3.74$  (i.e. good-reasonable). The  $C_{Da}$  model showed highest goodness-of-fit ( $x^2/df = 1.40$ ; good adjustment) and the SF the lowest one ( $x^2/df = 4.41$ ; reasonable adjustment).

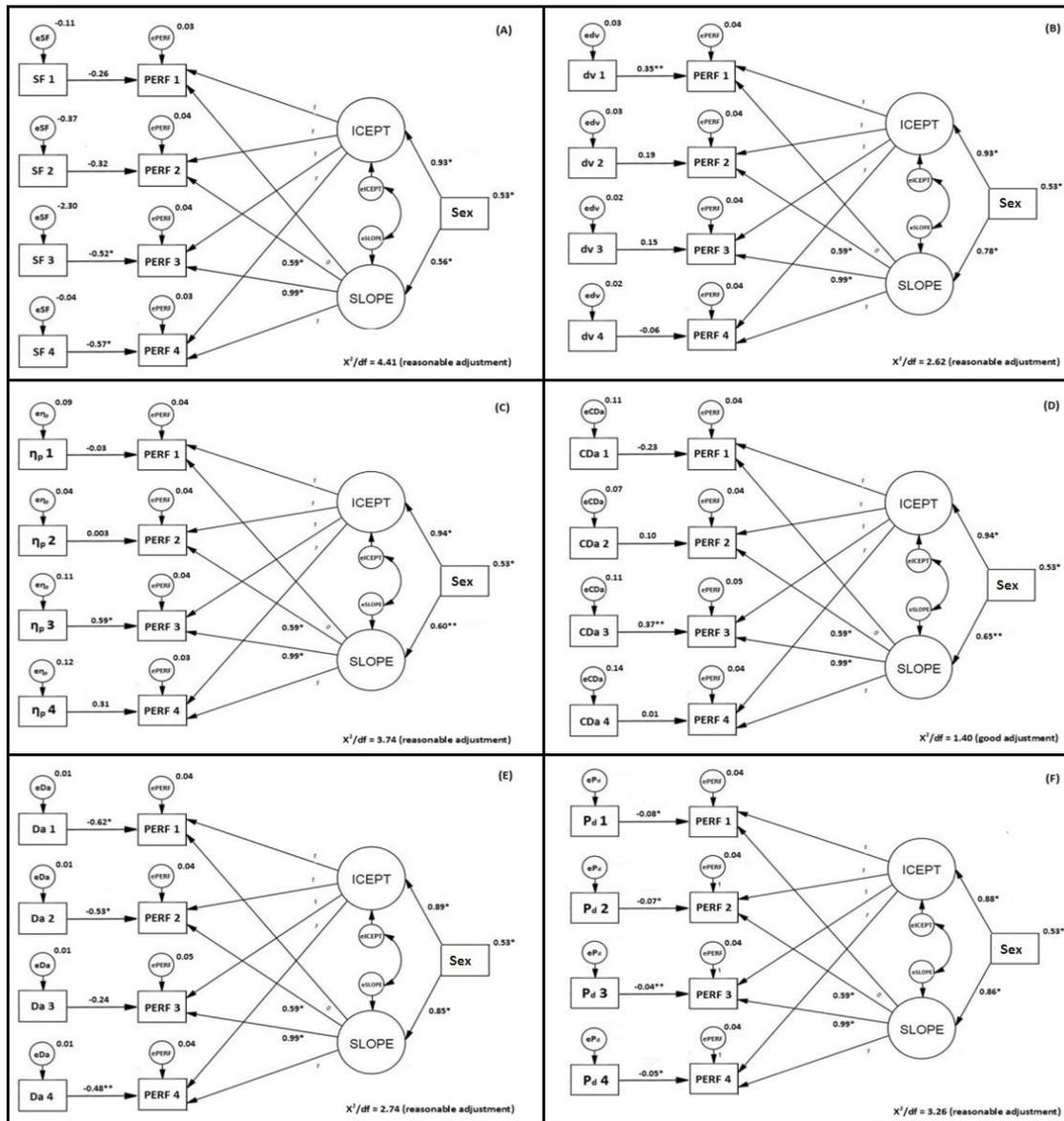


Figure 4. Growth confirmatory models for performance and effects of the selected variables. (A) SF – stroke frequency; (B) dv – intra-cyclic velocity fluctuation; (C)  $\eta_p$  – arm's propelling efficiency; (D) C<sub>Da</sub> – active drag coefficient; (E) D<sub>a</sub> – active drag; (F) P<sub>d</sub> – power to overcome drag; ICEPT – intercept effect; SLOPE – slope effect; Gender – gender effect; e<sub>xi</sub> – disturbance term for a given variable;  $x_i \rightarrow y_i$  – variable  $y_i$  depends from variable  $x_i$ .

### Discussion

The main aim of this study was to model a latent growth curve of swimming performance and its relationship with biomechanics over time to gather insight about the partial contribution of each factor and the gender effect. In the first two moments,

hydrodynamics was the major contributor to performance and in the last two, kinematics. The model was also able to detect a gender gap and a high intra- and inter-subject variability. Therefore, over a season, different determinant factors had a main influence on the performance enhancement for both boys and girls. Besides that, each one of them selected a unique strategy to enhance performance.

Cross-sectional studies showed that young swimmers' performance is highly influenced by kinematics and hydrodynamics (Morais et al., 2012). However, longitudinal follow-up studies that included these variables neglected the inter- and intra-subject changes (Latt et al., 2009). At least for adult swimming it was pointed out that intra-subject changes are not residual variance and it should not be disregarded in the overall analysis (Costa et al., 2013; Connaboy et al., 2010). The same idea was shared earlier by others for motor control (Komar et al., 2013) and kinematics changes (Figueiredo et al., 2012; Seifert et al., 2010). Latent growth curve modeling is able to estimate intra- and inter-subject variability. Variance analysis showed significant differences between swimmers at the baseline and during the performance growth. Residual variances tend to be neglected by other data analysis techniques (e.g. analysis of variance and multi-linear regressions). At least classical techniques are less sensitive to such residual variances. However, those variances are of major interest in latent growth curve modeling (Voelkle, 2007). A main finding of this research was that young swimmers presented a high intra- and inter-subject variability suggesting that each one has a very unique strategy to excel.

Latent growth curve modeling provides the amount of performance that is achieved in intermediate moments. Between M1 and M2 performance reached 59% of its final value in M4. Between competitive seasons, young swimmers have a break period impairing their energetics and kinematics (Moreira et al., 2014). The improvement between M1 and M2 might be related with the first meso-cycle that is characterized by a fairly high volume after the summer break (Figure 1). Afterwards, performance improved 39% (between M2 and M3) and 1% (between M3 and M4). Hence, as the major competition of the season is approaching, improvements are less sharp and meaningful. Similar trend is reported for adult/elite swimmers. Building-up for the major competition, adult swimmers are getting closer from their reserve upper-limits, and it is more challenging for any further improvement (Costa et al., 2013).

A gender gap was also identified at the baseline and during the performance growth. There is a very solid body of knowledge about the gender differences for pre- and post-pubertal athletes (Seifert et al., 2010). Literature reports that boys have a higher  $d_v$ ,  $D_a$ , and SF than girls (Barbosa et al., 2010). Therefore longitudinal structural equation modeling was successful in identifying the well-known gender gap. In this sense, the technique used is also sensitive enough whenever pooled data (both genders) is computed.

Swimming is characterized by the multi-dimensional interplay of different variables that will influence the performance. One might claim that the partial contribution of each exogenous variable to the endogenous one will change over time. That is, the partial contribution of each variable will not be constant over time. However, until now as much as we are aware no paper reported or quantified such phenomenon. Structural equation modeling is very sensitive to such changes and can be used to learn about it. In M1 and M2,  $D_a$  was the main performance determinant. Between M1 and M2 periodization included the highest volume of the season (Figure 1). The goals of those meso-cycles were to build-up energetics (mainly aerobic capacity) and improve technique. It was reported that training based on technical drills and kinesthetic feedbacks improved young swimmers' hydrodynamics and performance (Havriluk, 2006).  $D_a$  is strongly related to swimming velocity (equation 3). Hence, the increase in speed and therefore in performance lead to a higher  $D_a$ .  $C_{D_a}$  had a minor influence on performance growth. So, it can be speculated that the performance enhancement during this time frame might be more related to energetic build-up and less to technique enhancement.

In M3,  $\eta_p$  had the highest direct effect on the performance. Between M2 and M3 periodization was characterized by a decrease in total volume (Figure 1). These meso-cycles were more focused on technical parameters (enhancing stroke mechanics). This explains why on average the swimmers achieved the highest  $\eta_p$  in M3. Since long there has been a discussion whether young swimmers training should rely more on energetics or efficiency. Cross-sectional confirmatory models suggested that 50-60% of performance in these age-groups is related to biomechanics and technique enhancement (Morais et al., 2012). In M4, SF was the variable presenting the highest direct effect on performance. Between M3 and M4 periodization included

an increase in the aerobic power and aerobic capacity sets (Figure 1). This was coupled with a slight increase of the dry-land training sessions that included strength power routines. For adolescent sprint swimmers, an association was found between high muscular strength parameters and an increase in SF (Girould et al., 2007). At least in adult swimmers aerobic power paces are related to customize SF-stroke length relationships (McClean et al., 2010; Wakayoshi et al., 1995). Therefore, to swim at aerobic power sets a fairly constant and high stroke length with a high SF is needed. To be able to optimize this SF-stroke length relationship dry-land power training is a must.

It was attempted in one of the earliest models to include anthropometrics variables to control the potential confounding factor of the maturation and growth. However, after running the model, we failed to obtain significant results and a reasonable adjustment. Because we track down and follow-up subjects in Tanner 1-2, one might consider that most of them are yet pre-pubescent and therefore one single year is not enough to verify significant changes in biological maturation. With this we are not suggesting that they are not in a process of biological development but only that because they did not reach any spur, it is more challenging to have anthropometrics as a determinant factor. However, later on, that is, swimmers in the following Tanner stages this is more obvious (Jurimae et al., 2007; Falk et al., 2004). Overall, in M1 and M2 hydrodynamics (i.e.  $D_a$ ) was the major contributor to performance while in M3 and M4 was the kinematics ( $\eta_p$  and SF, respectively). Therefore, the main determinant at a given moment is related to the periodization model designed. It is possible to design models that are more based on energetics (M1 and M2) or technique (M3 and M4). A model that relies more on energetics allows a very quick and sharp improvement, but on the other hand the efficiency is compromised and increases the odds of an early burn-out. A model that is based on the technique is more time-consuming and performance enhancement might take some time to happen. However, a proper technique will be needed for further improvement reaching adulthood, when most of the periodization is energetically oriented (e.g. Schnitzler et al., 2014). Besides, it is at these early ages that the motor learning mechanisms of any skill is more effective. Considering the pros and cons of each approach, an age-group coach should consider to compromise both (energetics build-up and technique enhancement) but putting more focus on the technique and

efficiency if the athlete's career is to be seen in the long-run. Hence, it seems that many of the changes in performance can be attributed to the type of training that swimmers were undergoing at the time of each data collection. This could be useful for coaches as it shows that technical parameters are the most determinant ones in the young swimmers' performance improvement. They can apply these technical drills according to their macro-cycles, avoiding the athletes to burn out with high amounts of training workloads, especially close to the main events.

## Conclusion

Latent modeling is a comprehensive way to gather insight about young swimmers' performance over time. This was showcased with young swimmers engaged in a national talent ID scheme. Different variables were responsible for the performance improvement over the season. A significant intra- and inter-subject variability was verified. These findings suggest that a very unique and customized strategy is used by each swimmer to excel. Overall it seems that young swimmers' coaches should put the focus on the hydrodynamic profile and also on the stroke mechanics (i.e. technical ability) to enhance the performance, notably sprinters. Moreover the performance main determinants are also related to the training periodization.

## References

- Barbosa TM, Bragada JA, Reis VM, Marinho DA, Carvalho C, Silva AJ. (2010). Energetics and biomechanics as determining factors of swimming performance: updating the state of the art. *J Science Med Sport*. 13: 262-269.
- Bideault G, Herault R, Seifert L. (2013). Data modelling reveals inter-individual variability of front crawl swimming. *J Science Med Sport*. 16: 281-285.
- Biesanz JC, West SG, Kwok OM. (2003). Personality over time: Methodological approaches to the study of short-term and long-term development and change. *J Pers*. 71: 905–941.
- Castellanos-Ryan N, Parent S, Vitaro F, Tremblay RE, Séquin JR. (2013). Pubertal development, personality, and substance use: a 10 year longitudinal study from childhood to adolescence. *J Abnorm Psychol*. 122: 782-796.

Connaboy C, Coleman S, Moir G, Sanders R. (2010). Measures of reliability in the kinematics of maximal undulatory underwater swimming. *Med Sci Sports Exerc.* 42: 762-770.

Costa MJ, Bragada JA, Meijas JE, Louro H, Marinho DA, Silva AJ, Barbosa TM. (2013). Effects of swim training on energetics and performance. *Int J Sports Med.* 34: 507-513.

Falk B, Bronshtein Z, Zigel L, Constantini N, Eliakim A. (2004). Higher tibial quantitative ultrasound in young female swimmers. *Br J Sports Med.* 38: 461-465.

Figueiredo P, Seifert L, Vilas-Boas JP, Fernandes RJ. (2012). Individual profiles of spatio-temporal coordination in high intensity swimming. *Hum Mov Sci.* 31: 1200-1212.

Girould S, Maurin D, Dugué B, Chatard JC, Millet G. (2007). Effects of dry-land vs. resisted- and assisted-sprint exercises on swimming sprint performances. *J Strength Cond Res.* 21: 599-605.

Havriluk R. (2006). Magnitude of the effect of an instructional intervention on swimming technique and performance. In: JP Vilas-Boas, F Alves, A Marques (Eds), *X International Symposium of Biomechanics and Medicine in Swimming* (pp. 218-220). Porto: Portuguese Journal of Sport Sciences.

Jurimae J, Haljaste K, Cicchella A, Latt E, Purge P, Leppik A, Jurimae T. (2007). Analysis of swimming performance from physical, physiological, and biomechanical parameters in young swimmers. *Pediatr Exerc Sci.* 19: 70-81.

Kolmogorov S, Duplisheva, O. (1992). Active drag, useful mechanical power output and hydrodynamic force in different swimming strokes at maximal velocity. *J Biomech.* 25: 311-318.

Komar J, Sanders RH, Chollet D, Seifert L. (2013). Do qualitative changes in inter-limb coordination lead to effectiveness of aquatic locomotion rather than efficiency? *J Appl Biomech.* 30:189-196.

Latt E, Jürimäe J, Haljaste K, Cicchella A, Purge P, Jurimae, T. (2009). Longitudinal development of physical and performance parameters during biological maturation of young male swimmers. *Percept Mot Skills.* 108: 297-307.

Maia JA, Beunen G, Lefevre J, Claessens AL, Renson R, Vanreusel B. (2003). Modeling stability and change in strength development: a study in adolescent boys. *Am J Hum Biol.* 15: 579-591.

Marinho DA, Barbosa TM, Costa MJ, Figueiredo C, Reis VM, Silva AJ, Marques MC. (2010). Can 8 weeks of training affect active drag in young swimmers? *J Sports Sci Med.* 9: 71-78.

McLean SP, Palmer D, Ice G, Truijens M, Smith JC. (2010). Oxygen uptake response to stroke rate manipulation in freestyle swimming. *Med Sci Sports Exerc.* 42: 1909-1913.

Morais JM, Jesus S, Lopes V, Garrido N, Silva AJ, Marinho DA, Barbosa TM. (2012). Linking selected kinematic, anthropometric and hydrodynamic variables to young swimmer performance. *Pediatr Exerc Sci.* 24: 649-664.

Moreira M, Morais JE, Marinho DA, Silva AJ, Barbosa TM, Costa MJ. (2014). Growth influences biomechanical profile of talented swimmers during the summer break. *Sports Biomech.* 13: 62-74.

Park I, Schutz RW. (2005). An introduction to latent growth models: analysis of repeated measures physical performance data. *Res Quart Exerc Sport.* 76: 176-192.

Schnitzler C, Seifert L, Chollet D, Toussaint H. (2014). Effect of aerobic training on inter-arm coordination in highly trained swimmers. *Hum Mov Sci.* 33: 43-54.

Seifert L, Leblanc H, Herault R, Komar J, Button C, Chollet D. (2011). Inter-individual variability in the upper-lower limb breaststroke coordination. *Hum Mov Sci.* 30: 550-565.

Seifert L, Leblanc H, Chollet D, Delignières D. (2010). Inter-limb coordination in swimming: effect of speed and skill level. *Hum Mov Sci.* 29: 103-113.

Seifert L, Barbosa TM, Kjendlie PL. (2010). Biophysical approach to swimming: gender effect. In SA Davies (Ed), *Gender Gap: Causes, Experiences and Effects*, (pp. 59-80). New York: Nova Science Publishers.

Silva AF, Figueiredo P, Seifert L, Soares S, Vilas-Boas JP, Fernandes RJ. (2013). Backstroke technical characterization of 11-13 year old swimmers. *J Sports Sci Med.* 12: 623-629.

Tanner JM. (1962). *Growth at adolescence.* (2<sup>nd</sup> ed). Oxford: Blackwell Scientific.

Toubekis AG, Vasilaki A, Douda H, Gourgoulis V, Tokmakidis S. (2011). Physiological responses during interval training at relative to critical velocity intensity in young swimmers. *J Sci Med Sport.* 14: 363-368.

Voelkle MC. (2007). Latent growth curve modeling as an integrative approach to the analysis of change. *Psychol Sci.* 49: 375-414.

Wakayoshi K, D'Acquisito J, Cappert JM, Troup JP. (1995). Relationship between oxygen uptake, stroke rate and swimming velocity in competitive swimming. *Int J Sports Med.* 16: 19-23.

Wheaton B. (1987). Assessment of fit in overidentified models with latent variables. *Sociol Methods Res.* 16: 118-154.

Wu W, Taylor AB, West SG. (2009). Evaluating model fit for growth curve models: integration of fit indices from SEM and MLM frameworks. *Psychol Methods.* 14 : 183-201.

Zamparo P, Pendergast DR, Mollendorf J, Termin A, Minetti, A. E. (2005). An energy balance of front crawl. *Eur J Appl Physiol.* 94: 134–144.

---

## **Chapter 3**

**Cluster stability as a new method to assess changes in performance and its determinant factors over a season in young swimmers**

---



## Abstract

The purpose of this study was to (i) apply a new method to identify, classify and follow-up young swimmers, based on their performance and its determinant factors over a season and (ii) analyze the swimmers' stability over a competitive season with that method. Fifteen boys and eighteen girls ( $11.8 \pm 0.7$  years) part of a national talent ID scheme were evaluated at three different moments of a competitive season. Performance (i.e. official 100-m freestyle race time), arm span, chest perimeter, stroke length, swimming velocity, intra-cyclic velocity fluctuation, coefficient of active drag, propelling efficiency and stroke index were selected as variables. Hierarchical and k-Means cluster analysis were computed. Data suggested a three cluster solution, splitting the swimmers according to their performance in all three moments. Cluster 1 was related to better performances ("talented" swimmers), cluster 2 to poor performances ("no-proficient" swimmers) and cluster 3 to average performance ("proficient" swimmers) in all moments. Stepwise discriminant analysis revealed that 100%, 94% and 85% of original groups were correctly classified for the first, second and third evaluation moments, respectively ( $0.11 \leq \Lambda \leq 0.80$ ;  $5.64 \leq X^2 \leq 63.40$ ;  $0.001 < P \leq 0.68$ ). Membership of clusters was moderately stable over the season (stability range 46.1% to 75% for the two clusters with most subjects). Cluster stability is a feasible, comprehensive and informative method to gain insight into changes in performance and its determinant factors in young swimmers. "Talented" swimmers were characterized by anthropometrics and kinematic features.

**Key-words:** *prepubescent swimmers, seasonal adaptations, longitudinal assessment, classification*



## Introduction

Two of the most interesting research topics in the field of sports performance, and specifically in competitive swimming, are the identification of performance determinant factors and the performance modeling. Several research groups focused on identifying the performance determinants, such as the following: (i) what are the main determinants and (ii) how do they interplay to improve performance. Performance of young swimmers is influenced by growth and maturation (Morais et al., 2013). Biological maturation may promote changes in their biomechanics, motor control, energetics, which may affect their expertise achievement (Latt et al., 2009). Young swimmers are exposed to different rates of development which progress according to their own time scale (Liu et al., 2006). E.g. a couple of structural equation models reported that anthropometrics influences swimmers' kinematics and hence their performance (Morais et al., 2012; Barbosa et al., 2010a). The second topic of research interest is to model performance over time. The model enables a researcher or sports analyst to predict one's performance at a given moment, e.g. say at a given age or competition (i.e. mean stability, within-subjects analysis) (Hopkins and Hewson, 2001). Longitudinal assessments can also be carried out to understand the relative changes of performance among the main athletes (i.e. normative stability, between-subjects analysis) (Morais et al., 2013; Costa et al., 2011).

New trends in sports performance and expertise should adopt a multi-disciplinary approach to enhance the understanding of the athlete-environment relationship, as exemplifying a complex and dynamic system in opposition to the traditional frameworks (Phillips et al., 2010). In such a dynamic system, all expert athletes do not follow the same pathway to achieve a given performance (Komar et al., 2014; Durand-Bush and Salmela, 2002). Keeping in view the complex and diverse nature of the scientific fields that play a role in performance, and despite the existence of an optimal pathway to expertise achievement, each athlete selects a customized path (Durand-Bush and Salmela, 2002). Likewise, the performance of both adult/elite (Seifert et al., 2014; Barbosa et al., 2010b; Seifert et al., 2009) and young swimmers (Barbosa et al., 2010a) is determined by several domains. Probably, the partial contribution of each domain or determinant factor to performance will change over

time, and not remain constant. However, until now, no research was conducted on this aspect, at least in respect of young athletes who are typically involved in such complex and dynamic system, as has been done in the case of age-group swimmers.

Longitudinal and multi-disciplinary designs should help in understanding the performance changes and the partial contribution of each determinant factor over time in young swimmers. For an insight into this problem, two independent procedures must be selected (deterministic analysis and longitudinal/stability analysis). Identified with the deterministic analysis are the main factors acting upon the swimmer's performance at a given moment; in longitudinal analysis, changes of each selected variable are tracked down and followed-up. As these two procedures are independent, it is challenging to establish any causality among them. For a deeper insight of the relationship between these two analyses, it is worthwhile to merge the two into one single procedure. By doing so, it would be possible to learn: (i) about the changes in performance or determinant factors over time; (ii) how these factors interplay at a given moment and over time; and (iii) what are the main determinant factors at a given moment. This procedure can be applied for both short and long time-frames (e.g. from few weeks to several years) depending on the nature of the research.

As sports performance is a multi-disciplinary phenomenon, multivariate data analysis (e.g. cluster analysis) can be implemented to detect patterns within high dimensional datasets. Cluster analysis is one such procedure that identifies homogeneous groups of subjects. Subjects grouped in a specific cluster share several common characteristics, but are very dissimilar to others not belonging to that cluster (Rein et al., 2010). This procedure has been mainly applied in scientific fields such as genetics (Handl et al., 2005), motor control (Rein et al., 2010; Chow et al., 2008) and psychology (Clatworthy et al., 2005). A few studies have been conducted on adult/elite swimming to classify the coordination patterns (Bideault et al., 2013), start patterns (Vantorre et al., 2010) and race analysis (Chen et al., 2008). Cluster analysis can also be a feasible approach to identify and classify young athletes' determinant performance factors at a given moment. Over and above that, changes of a subject from one cluster to another, and in evaluation moments could enable one to understand the subject's stability and the reason behind the subject's performance

change. So, a new method that combines cluster analysis with longitudinal design should be tested. This novel method, based on cluster stability might provide us details about how determinant variables in isolation (i.e. bivariate analysis) or in combination (i.e. multivariate analysis) contribute to performance and how their partial contribution changes over time.

To our knowledge, no studies were carried out on cluster stability in swimming or even sports performance, excepting the one relating to nutrition. In that study, the authors developed a clustering solution about dietary patterns and thereafter analyzed the changes in the stability of young subjects between cluster groups in a given time-frame (Northstone et al., 2012). Hence, this method of assessing clusters stability over time can be considered a breakthrough in sport sciences, notably in sports performance and swimming. By performing cluster analysis, it will be possible not only to classify young swimmers according to their performance and its main determinants, but also to assess the stability of the cluster membership over time and thereby help coaches in following-up the athletes and designing customized trainings.

The aim of this study was: (i) to apply a new method to identify, classify and follow-up young swimmers, based on their performance and its determinant factors over a season and (ii) to analyze the swimmers' stability over a competitive season with that method. It was hypothesized that: (i) the new method is feasible and informative to identify, classify and follow-up young swimmers and; (ii) there is a moderate-to-high stability in the cluster membership across the season.

## **Methods**

### **Subjects**

Thirty-three young swimmers (overall:  $11.8 \pm 0.7$  years,  $262.6 \pm 74.3$  FINA points at SCM 100-m freestyle,  $47.0 \pm 8.3$  kg body mass,  $156.2 \pm 8.8$  cm height; boys:  $N = 15$ ,  $12.3 \pm 0.6$  years,  $227.9 \pm 69.8$  FINA points at SCM 100-m freestyle,  $49.9 \pm 9.3$  kg body mass,  $159.9 \pm 8.7$  cm height; girls:  $N = 18$ ,  $11.7 \pm 0.9$  years,  $291.1 \pm 66.2$  FINA points at SCM 100-m freestyle,  $44.6 \pm 6.7$  kg body mass,  $153.1 \pm 7.8$  cm height; Tanner stages 1-2 by self-report) participating on a regular basis in regional and national level competitions were evaluated. The sample included age-group national

record holders, age-group national champions and other swimmers who formed part of a national talent ID scheme. At the baseline, the swimmers had  $3.18 \pm 0.52$  years of training experience. The swimmers had  $5.59 \pm 0.92$  (ranging from 3 to 8 in the season; 90 min each session) weekly training sessions, including warm-up, recovery, slow, medium and intense pace, technical drills, as well as dry-land strength and conditioning sessions (2 per week).

Coaches and/or parents and also the swimmers gave their consent for the swimmers participation on this study. All procedures were in accordance to the Helsinki Declaration regarding Human research. The University of Trás-os-Montes and Alto Douro Ethic Committee also approved the study design (ethic review: UTAD-2011-219).

### **Study design**

A longitudinal research design of selected variables over three different moments of the season was carried out. The swimmers were evaluated first in (i) October (M1, the season's first competition), (ii) March (M2, the winter's main competition) and; (iii) June (M3, the summer's main competition). Variables that are regularly reported as having an effect on swimming performance (Seifert et al., 2014; Komar et al., 2014; Morais et al., 2013; Morais et al., 2012; Barbosa et al., 2010a; Barbosa et al., 2010b; Seifert et al., 2009) were selected. All pool testing data collection was conducted with no swimmers in nearby lanes to reduce drafting or pacing effects.

### **Performance data collection**

For assessing swimming performance, the 100-m freestyle race time, recorded officially at regional or national short course meter swimming pool (i.e. 25-m length) was selected. The time gap between data collection and the race was no more than two weeks (Barbosa et al., 2010a).

### **Anthropometric data collection**

Arm span (AS) was measured standing in the upright position with arms and fingers fully extended and abducted at  $90^\circ$ . The distance between the third fingertip of each hand was measured with a flexible anthropometric tape (RossCraft, Canada) (ICC = 0.98). Chest perimeter (CP), was measured with a flexible anthropometric tape

(RossCraft, Canada) when the swimmer simulated a streamlined gliding (i.e. hydrodynamic) position with both arms fully extended upwards (ICC = 0.99).

### **Kinematic data collection**

Swimming velocity ( $v$ ), stroke length (SL) and intra-cyclic velocity fluctuation ( $dv$ ) were selected as kinematic variables. Swimmers performed a standardized warm-up of approximately 1,000m (Neiva et al., 2014). Afterwards, each swimmer performed three maximal 25-m trials in freestyle with push-off start. Swimmers were advised to reduce gliding during the push-off. Between each trial, the swimmers were allowed 30 minute rest to ensure full recovery. For further analysis, the average value of three trials was considered (ICC = 0.96).

A speedo-meter cable (Swim speedo-meter, Swimsportec, Hildesheim, Germany) was attached to the swimmers' hip. A 12-bit resolution acquisition card (USB-6008, National Instruments, Austin, Texas, USA) was used to transfer data ( $f = 50$  Hz) from the speedo-meter to a software interface in LabVIEW® (v.2009) (Barbosa et al., 2013). Data was exported to a signal processing software (AcqKnowledge v. 3.5, Biopac Systems, Santa Barbara, USA) and filtered with a 5Hz cut-off low-pass 4<sup>th</sup> order Butterworth filter.

Swimming velocity (in  $\text{m}\cdot\text{s}^{-1}$ ) was calculated as the time spent between the 5<sup>th</sup> and 20<sup>th</sup> meter (i.e. middle 15-m;  $v=d/t$ ). Stroke length (in m) was calculated by dividing velocity with stroke frequency ( $SL = v/SF$ ) (Craig and Pendergast, 1979). Stroke frequency was measured (in  $\text{cycles}\cdot\text{min}^{-1}$  and then converted to Hz) by two expert evaluators with a stroke counter (base 3). Intra-cyclic velocity fluctuation (in dimensionless units) was calculated with the coefficient of variation as reported elsewhere ( $dv = CV = \text{standard deviation}/\text{mean}$ ) (Barbosa et al., 2010b).

### **Hydrodynamic data collection**

Coefficient of active drag ( $C_{Da}$ ) was computed using the velocity perturbation method (Kolmogorov and Duplisheva, 1992). To calculate  $C_{Da}$ , the following inputs are required: water density (being  $1000 \text{ kg}/\text{m}^3$ ), active drag force (given by the difference in velocity swimming with and without perturbation buoy according to its resistance,  $Da = [(Db \cdot vb \cdot v^2)/(v^3 - vb^3)]$ ), swimming velocity and the swimmer's projected frontal surface area (Kolmogorov and Duplisheva, 1992). Each swimmer performed two

extra maximal 25-m freestyle trials in freestyle with a push-off start. The first trial was performed without the perturbation device and the second one with the perturbation device (Marinho et al., 2010).

Swimming velocity was assessed between the 11<sup>th</sup>-m and 24<sup>th</sup>-m from the starting wall (Marinho et al., 2010). The time spent to cover this distance was measured with a manual stopwatch (Golfinho Sports MC 815, Aveiro, Portugal) by two expert evaluators (ICC = 0.96). The evaluators followed the swimmer to have a good line of sight when the swimmer passed the two distance marks.

The swimmers' projected frontal surface area was measured using a photogrammetric technique (Morais et al., 2011), and their photographs taken with a digital camera (DSC-T7, Sony, Tokyo, Japan) in the transverse plane from above. While taking photographs, the swimmers stood on land, in the upright and streamlined position. In this position, the arms were fully extended above the head, one hand over the other, and the fingers also extended close together while the head was in neutral position. They wore a regular textile swim suit, cap and goggles. On the camera shooting field, a calibration frame with 0.945-m length was beside the swimmer at the shoulder level. The S was measured with an area measuring software (Udruler, AVPSOft, USA) after importing the digital picture (ICC = 0.97).

### **Efficiency data collection**

Efficiency variables were estimated from kinematic data. Stroke index (SI, in  $\text{m}^2 \cdot \text{s}^{-1}$ ) was calculated as the product of SL and  $v$  ( $\text{SI} = \text{SL} \cdot v$ ) (Costill et al., 1985). The arm's propelling efficiency ( $\eta_p$ , in %) was also calculated, using  $v$ , SF and the distance between the shoulder and the tip of the 3<sup>rd</sup> finger during the insweep (in m) as inputs (Zamparo et al., 2005). The shoulder-finger distance was computed trigonometrically measuring the arm's length and considering the average elbow angles during the insweep of the arm pull (Zamparo et al., 2006).

### **Statistical analysis**

Kolmogorov-Smirnov and Levene tests were used to analyze normality and homocedasticity assumptions, respectively. Mean, one standard-deviation, minimum and maximum were calculated as descriptive statistics.

To increase confidence in the stability of the emergent profiles, two clustering approaches were used: (i) hierarchical cluster analysis (using Ward's linkage method with squared Euclidian distance measure to provide guidance as to the number of clusters represented in the data); (ii) k-Means (non-hierarchical) cluster analysis to compute the clusters and thus group the swimmers according to their similarities. K-means defines a prototype in terms of a centroid (i.e. the mean of a group of points), typically applied to objects in a continuous n-dimensional space. Standardized z-scores of the selected variables were calculated to compare datasets with different units and/or magnitudes (Rein et al., 2010).

ANOVA was used to identify the variables having the highest influence in each cluster and discriminant analysis (stepwise method) was used to validate them ( $P \leq 0.05$ ). Total eta square ( $\eta^2$ ) was selected as effect size index and interpreted as (Ferguson, 2009): (i) without effect if  $0 < \eta^2 \leq 0.04$ ; (ii) minimum if  $0.04 < \eta^2 \leq 0.25$ ; (iii) moderate if  $0.25 < \eta^2 \leq 0.64$  and; (iv) strong if  $\eta^2 > 0.64$ . Swimmers' changes between clusters were assessed by cross-tabulating cluster solutions at different moments. This kind of assessment enables calculating the proportion of swimmers remaining in the same cluster between each pair of moments and consequently the proportion of swimmers that changed between clusters (Northstone et al., 2012). The distance between clusters enables to learn about clusters' similarities/dissimilarities. A higher distance between clusters means a higher dissimilarity.

## Results

Overall, the anthropometric features increased between the first and last evaluation moments (body mass increased from  $49.9 \pm 9.3$  to  $52.9 \pm 9.1$  kg in case of boys, and from  $44.6 \pm 6.7$  to  $46.5 \pm 6.2$  kg in case of girls; height increased from  $159.9 \pm 8.7$  to  $162.9 \pm 8.8$  cm in case of boys, and from  $153.1 \pm 7.8$  to  $155.6 \pm 7.2$  cm in case of girls). Coefficient of determination ( $R^2$ ) was selected to test several cluster solutions (from 1 to 9, i.e.  $1 < k < 9$ ). A three-cluster solution ( $k = 3$ ) provided stable interpretations over the season.

The SI, v and SL, were the variables with the strongest (i.e.  $\eta^2 > 0.64$ ) and the best discrimination effect (i.e. highest F-ratios) among the cluster solutions of all moments (Table 1). Cluster 1 was characterized by high CP, AS and SI (M1), CP and AS (M2),

AS, SI and  $v$  (M3). Cluster 1 was also characterized by the best performance at all moments (labeled as “talented” swimmers; i.e. better performing swimmers). Cluster 2 was characterized by high  $dv$  (M1 and M2) and  $\eta_p$  (M3). Cluster 2 was also characterized by the slowest performance at all moments (labeled as “no-proficient” swimmers; i.e. poorer performing swimmers). Cluster 3 was characterized by high  $CD_a$  (M1), SI,  $\eta_p$  and  $v$  (M2), CP and AS (M3). Cluster 3 was also characterized by average performance at all moments (labeled as “proficient” swimmers; i.e. average performing swimmers).

Table 1. Performance and its determining factors for swimmers classified in each cluster assessed at baseline (moment 1), mid-season (moment 2), and end-season (moment 3).

	Moment 1									
	Cluster 1 (n = 13)		Cluster 2 (n = 8)		Cluster 3 (n = 12)		F	P	$\eta^2$	
	Mean $\pm$ 1SD	z	Mean $\pm$ 1SD	z	Mean $\pm$ 1SD	z				
AS [cm]	166.9 $\pm$ 9.6	0.81	149.8 $\pm$ 9.3	-0.73	153.6 $\pm$ 5.4	-0.39	13.1	<0.001	0.40	
CP [cm]	83.6 $\pm$ 3.4	0.92	72.8 $\pm$ 4.5	-0.91	75.9 $\pm$ 3.4	-0.38	23.8	<0.001	0.50	
SL [m]	1.69 $\pm$ 0.17	0.80	1.13 $\pm$ 0.20	-1.24	1.46 $\pm$ 0.11	-0.04	28.0	<0.001	0.66	
v [ $m \cdot s^{-1}$ ]	1.38 $\pm$ 0.11	0.80	0.91 $\pm$ 0.16	-1.34	1.21 $\pm$ 0.08	0.02	38.5	<0.001	0.76	
dv [dimensionless]	0.09 $\pm$ 0.03	0.03	0.10 $\pm$ 0.04	0.29	0.08 $\pm$ 0.02	-0.23	0.6	0.517	0.10	
C <sub>Da</sub> [dimensionless]	0.31 $\pm$ 0.08	-0.19	0.23 $\pm$ 0.08	-0.68	0.45 $\pm$ 0.20	0.67	6.5	0.004	0.38	
$\eta_p$ [%]	30.47 $\pm$ 2.93	0.63	22.46 $\pm$ 4.39	-1.14	27.97 $\pm$ 2.86	0.07	14.6	<0.001	0.57	
SI [ $m^2 \cdot s^{-1}$ ]	2.36 $\pm$ 0.36	0.88	1.06 $\pm$ 0.32	-1.30	1.78 $\pm$ 0.22	-0.09	43.6	<0.001	0.75	
PERF [s]	71.17 $\pm$ 5.91	-0.75	83.67 $\pm$ 5.11	1.00	77.57 $\pm$ 4.44	0.14	14.5	<0.001	0.49	
	Moment 2									
	Cluster 1 (n = 8)		Cluster 2 (n = 17)		Cluster 3 (n = 8)		F	P	$\eta^2$	
	Mean $\pm$ 1SD	Z	Mean $\pm$ 1SD	z	Mean $\pm$ 1SD	z				
AS [cm]	173.1 $\pm$ 9.2	1.20	155.1 $\pm$ 6.4	-0.52	159.7 $\pm$ 8.1	-0.08	15.2	<0.001	0.77	
CP [cm]	86.5 $\pm$ 3.6	0.97	77.5 $\pm$ 4.5	-0.47	80.6 $\pm$ 6.4	0.03	13.1	<0.001	0.40	
SL [m]	1.30 $\pm$ 0.17	0.66	1.04 $\pm$ 0.10	-0.64	1.30 $\pm$ 0.19	0.70	43.9	<0.001	0.49	
v [ $m \cdot s^{-1}$ ]	1.16 $\pm$ 0.23	0.87	0.85 $\pm$ 0.06	-0.66	1.10 $\pm$ 0.16	0.54	74.4	<0.001	0.53	
dv [dimensionless]	0.11 $\pm$ 0.05	0.32	0.10 $\pm$ 0.03	-0.01	0.10 $\pm$ 0.02	-0.29	0.1	0.918	0.07	
C <sub>Da</sub> [dimensionless]	0.302 $\pm$ 0.08	0.16	0.26 $\pm$ 0.09	-0.31	0.30 $\pm$ 0.11	0.50	2.9	0.066	0.14	
$\eta_p$ [%]	22.71 $\pm$ 3.88	0.20	20.03 $\pm$ 2.12	-0.46	21.88 $\pm$ 4.02	0.77	39.3	<0.001	0.29	
SI [ $m^2 \cdot s^{-1}$ ]	1.55 $\pm$ 0.50	0.80	0.89 $\pm$ 0.13	-0.63	1.46 $\pm$ 0.40	0.55	72.0	<0.001	0.51	
PERF [s]	64.72 $\pm$ 4.88	-1.23	75.91 $\pm$ 3.98	0.61	72.18 $\pm$ 6.02	-0.07	13.4	<0.001	0.61	
	Moment 3									
	Cluster 1 (n = 6)		Cluster 2 (n = 18)		Cluster 3 (n = 9)		F	P	$\eta^2$	
	Mean $\pm$ 1SD	z	Mean $\pm$ 1SD	z	Mean $\pm$ 1SD	z				
AS [cm]	176.6 $\pm$ 7.4	1.41	156.6 $\pm$ 7.1	-0.55	162.3 $\pm$ 10.1	0.17	14.7	<0.001	0.61	
CP [cm]	89.2 $\pm$ 3.3	1.17	79.2 $\pm$ 5.76	-0.46	82.9 $\pm$ 6.1	0.14	8.4	0.001	0.43	
SL [m]	1.69 $\pm$ 0.15	1.17	1.34 $\pm$ 0.20	-0.36	1.41 $\pm$ 0.19	-0.05	25.2	<0.001	0.38	
v [ $m \cdot s^{-1}$ ]	1.49 $\pm$ 0.09	1.30	1.13 $\pm$ 0.16	-0.44	1.22 $\pm$ 0.18	0.02	58.3	<0.001	0.48	
dv [dimensionless]	0.10 $\pm$ 0.02	0.24	0.09 $\pm$ 0.02	0.03	0.08 $\pm$ 0.02	-0.23	0.1	0.953	0.05	
C <sub>Da</sub> [dimensionless]	0.52 $\pm$ 0.34	0.91	0.30 $\pm$ 0.12	-0.22	0.32 $\pm$ 0.09	-0.14	7.8	0.002	0.23	
$\eta_p$ [%]	29.13 $\pm$ 1.35	0.57	26.49 $\pm$ 3.91	-0.11	26.92 $\pm$ 3.83	-0.16	33.1	<0.001	0.09	
SI [ $m^2 \cdot s^{-1}$ ]	2.54 $\pm$ 0.29	1.39	1.54 $\pm$ 0.40	-0.44	1.76 $\pm$ 0.44	-0.03	64.1	<0.001	0.52	
PERF [s]	61.63 $\pm$ 2.90	-1.46	73.43 $\pm$ 3.92	0.60	68.64 $\pm$ 3.36	-0.23	9.4	0.001	0.13	

AS – arm span; CP – chest perimeter; SL – stroke length;  $v$  – swimming velocity;  $dv$  – intra-cyclic velocity fluctuation;  $C_{Da}$  – active drag coefficient;  $\eta_p$  – propelling efficiency; SI – stroke index; PERF – performance; SD – standard deviation;  $z$  – standardized data;  $F$  – F-ratio;  $P$  – significance value;  $\eta^2$  – effect size.

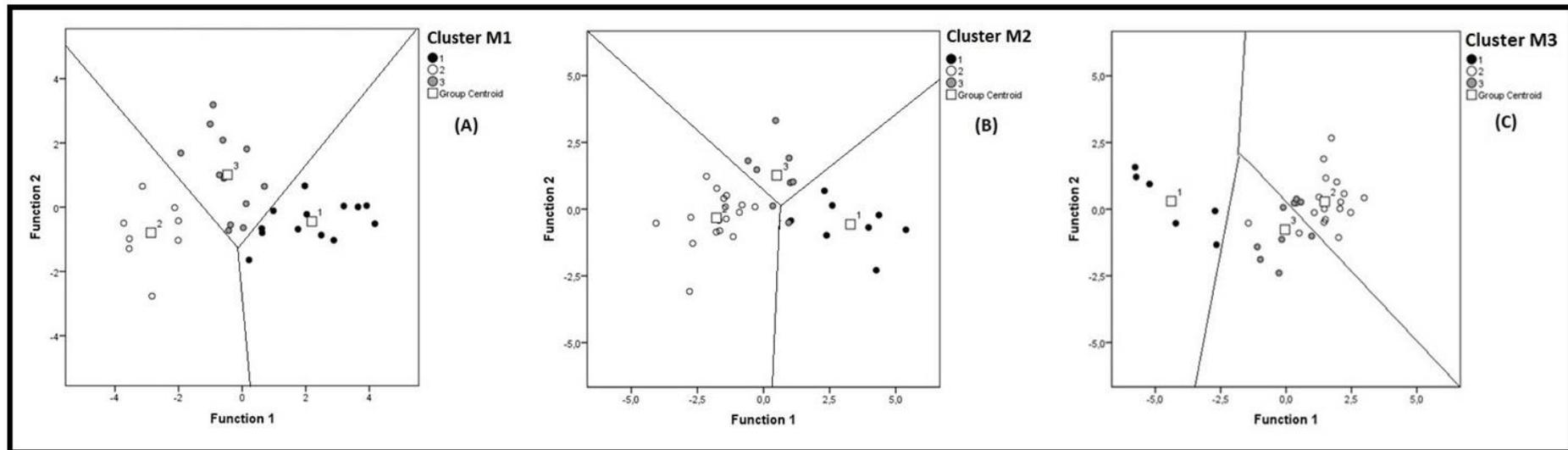


Figure 1. Territorial map of the two canonical discriminant functions in moment 1 (M1-A), moment 2 (M2-B) and moment 3 (M3-C), respectively. Group centroid 1 – “talented swimmers”; Group centroid 2 – “no-proficient swimmers”; Group centroid 3 – “proficient swimmers”.

A comparison of the classification results of original (i.e. the frequencies found in the data) and predicted (i.e. the predicted frequencies from the analysis) group memberships, according to the canonical discriminant functions obtained, and stepwise discriminant analysis reveal that 100%, 94% and 85% of the original groups were correctly classified in M1, M2 and M3, respectively ( $0.11 \leq \Lambda \leq 0.80$ ;  $5.64 \leq X^2 \leq 63.40$ ;  $0.001 < P \leq 0.68$ ). Visual inspection of the territorial map of the two canonical discriminant functions reveals a good or very good compactness and separation at M1, M2 and M3 (Figure 1).

As regards the cluster membership along the season (see Table 2), it can be seen that cluster 2 ("no-proficient" swimmers) had the highest stability (70.6% to 75% of the swimmers stayed in this cluster, at M2 vs M3 and M1 vs M2, as well as, M1 vs M3), followed by cluster 1 ("talented" swimmers) (46.1% at M1 vs M3 to 61.5% at M1 vs M2) and finally by cluster 3, which had the least stability ("proficient" swimmers) (from 0% at M2 vs M3 to 25% at M1 vs M2). Overall, there was thus a moderate stability in the clustering membership as the two clusters with more subjects presented a stability ranging roughly between 45 to 75%.

Table 2 also shows the distance between clusters' centers. At all moments, swimmers in cluster 3 ("proficient" swimmers) are closer to swimmers in cluster 1 ("talented" swimmers). However from M1 to M3, the cluster distance between cluster 3 ("proficient" swimmers) and cluster 2 ("no-proficient" swimmers) decreases from 3.091 to 1.851 (i.e. higher similarity). At the same time, the distance between cluster 3 ("proficient" swimmers) and cluster 1 ("talented" swimmers) increases from 2.728 to 4.474 (i.e. higher dissimilarity). This suggests that during this time-frame the "proficient" swimmers could not reach-up the "talented" swimmers, but at the time the "no-proficient" ones were able to close the gap to the "proficient" counterparts.

Table 2. Number of swimmers re-classified in each cluster between baseline (moment 1) and mid-season (moment 2), between mid-season (moment 2) and end-season (moment 3), and between baseline (moment 1) and end-season (moment 3). Distances between cluster centers for each pairwise comparison of clusters at each moment are also shown.

			Cluster 1		Cluster 2		Cluster 3	
			n	%	n	%	n	%
Cross-tabulations	M1 vs M2	Cluster 1	8	61.5	0	0	0	0
		Cluster 2	2	15.4	6	75	9	75
		Cluster 3	3	23.1	2	25	3	25
	M2 vs M3	Cluster 1	4	50	0	0	2	25
		Cluster 2	0	0	12	70.6	6	75
		Cluster 3	4	50	5	29.4	0	0
	M1 vs M3	Cluster 1	6	46.1	0	0	0	0
		Cluster 2	2	15.4	6	75	10	83.3
		Cluster 3	5	38.5	2	25	2	16.7
Distance between centers	M1	Cluster 1			5.101		2.728	
		Cluster 2	5.101				3.091	
		Cluster 3	2.728		3.091			
	M2	Cluster 1			5.076		3.330	
		Cluster 2	5.076				3.308	
		Cluster 3	3.330		3.308			
	M3	Cluster 1			5.870		4.474	
		Cluster 2	5.870				1.851	
		Cluster 3	4.474		1.851			

M1 – moment one; M2 – moment two; M3 – moment three; n – cluster total sample; % – cluster sample percentage.

## Discussion

The aim of this study was to describe and apply a new procedure to identify, classify and analyze the clusters' stability of young swimmers over a competitive season. The main finding was that cluster stability is a feasible, comprehensive and informative method to gain insight into young swimmers changes over time. Other important finding is that "talented swimmers" are characterized by anthropometrics and kinematic features.

The main goal of cluster analysis is to find similar trends within a dataset (young swimmers in this case). Participants or traits in the same cluster are similar to each other, while those in other clusters are as dissimilar as possible (Rein et al., 2010). A hierarchical model was used to define the number of clusters to retain with an  $R^2$  method, as suggested earlier (Vantorre et al., 2010). Afterwards, K-means solution was tested to compute the clusters and thereby grouped the swimmers according to

their similarities. Cluster analysis may be considered challenging because: (i) it might be difficult to form distinguished and equally dimensioned clusters (for this research, the sample sizes of the cluster solutions were fairly even) and; (ii) it may not consider the hypothetical relationships between variables (in this study, discriminant variables of the clustering solutions are meaningful for swimming researchers and practitioners). An interesting and novel idea is to combine cluster analysis with longitudinal analysis. This is based on the reasoning that after developing a cluster solution, it will analyze the membership changes over time (i.e. cluster stability). To the best of the authors' knowledge, this approach was not attempted so far in sports science. However, the present study proves it to be a feasible and informative way to gain insight into performance changes over time and the partial contribution of the determinant factors, or at least a set of factors, at a given moment.

Three clusters-solution ( $k = 3$ ) was the one that showed the highest power, besides allowing for a stable data interpretation. Marginal gains were observed after the 4<sup>th</sup> cluster ( $k \geq 4$ ). Cluster 1 was labeled as "talented" swimmers, cluster 2 as "no-proficient" swimmers and cluster 3 as "proficient" swimmers, because performance was the main discriminant variable across all the clusters at all moments. Cluster 2 was related to poor performances and high dv. There is evidence to show that a high dv is related to an increase in energy cost (Barbosa et al., 2010b). For the swimmers in Cluster 2, one might consider that their impaired performance is related to a high dv. Cluster 1 is related to better performance, anthropometrics (high AS and CP), and therefore to kinematics (high SI and v). Indeed, young swimmers' performance is highly related to anthropometric features (Komar et al., 2014). So, it seems that a few swimmers (from cluster 1) might rely more on their genetics and intrinsic characteristics (i.e. anthropometrics) than on external conditions (i.e. training and improving their technique) to enhance their performance. Cluster 3 was associated with a set of variables from different domains (i.e. anthropometric, kinematic, hydrodynamic and efficiency). For the swimmers of this cluster, the anthropometrics by themselves cannot explain their performance. Hence, another way to achieve better performance is to rely more on intervention programs (i.e. training sessions and technique improvements) than on the genetics.

The main novelty of this research was to assess cluster stability over time. Cluster analysis was developed to deal with problems in data mining when investigators needed to identify patterns in high-dimensional datasets (Rein et al., 2010), such as those associated with talent identification and follow-up. Discriminant analysis was used for clusters' validation (Milligan, 1981). Good-very good cluster separation and compactness are verified by visual inspection of the territorial map (Figure 1). Cross-validation is a comparison of the classification results of original data (i.e. the frequencies of groups found in the data) and those predicted (i.e. the frequencies of groups predicted from analysis) according to the canonical discriminant functions obtained. Cross-validation revealed that, along the season, the membership was correctly classified in 100%, 94% and 85% of the subjects (i.e. very good prediction). Cluster validity can also be assessed with bootstrapping (Seifert et al., 2011), normalized Hubert- $\Gamma$  (Rein, 2012). Cross-tabulation is a feasible and straightforward way to assess participants' changes between clusters across time (Northstone et al., 2012). The number of participants that remain, are added or removed from a cluster is calculated.

Across the three clusters, most outcomes between M1 and M3 showed improvement. A similar trend was reported by others for anthropometric, biomechanics and efficiency parameters (Morais et al., 2013; Latt et al., 2009). One might consider that improvement over time would happen in a linear or a near linear fashion-way in children. Surprisingly, although there was improvement in performance throughout the season (i.e. from M1 to M2 to M3), several determinant factors showed impairment. It seems that such non-linear changes were not reported so far for age-group swimmers. However, such changes were reported in the case of adult/elite counterparts (Costa et al., 2011). So, it seems that the determinant factors play a major role in contributing to performance at a given moment. This can be related to the designed periodization model, because age-group swimming, just as most youth sports, is designed with classic periodization models. Such models are based on one or two major peaks per season, one of them being the main competition. Therefore, coaches will be building-up fitness (i.e. energetics, and strength & conditioning) and improving techniques on the road to the main competition. Probably, because the main competition comes at the end of the season, coaches may consider that swimmers do not need to be in their best shape in the middle of the season, or at

least, they may rely more on a given set of determinant factors to improve their performance. Hence, near future research projects should consider to select a few energetic variables to control the role of the energetic build-up over the season in age-group swimmers, as happens on regular basis with adult/elite counterparts.

A moderate stability (i.e. moderate change in clusters' membership) was observed along the season. Cluster 1 ("talented" swimmers) presented a moderate stability (between 46.1% and 61.5%, even though the membership decreased from 13 to 6 swimmers). Cluster 3 ("proficient" swimmers") showed a low-moderate stability (between 0% and 25% and membership decrease from 12 to 9 swimmers). Overall, cluster membership of "talented" and "proficient" swimmers seems to have decreased along the season. Cluster 2 ("no-proficient" swimmers) presented a high stability (between 70.6% and 75% and membership increase from 8 to 18 swimmers). This increase is related to the movement (i.e. selection) of some swimmers from clusters 1 and 3 to cluster 2, because they could not maintain high performance levels. It should be noted that all swimmers improved their performances from M1 to M3 (i.e. within-subject comparison). Interestingly a couple of subjects moved straight from Cluster 1 to Cluster 2 between two evaluations moments. Such events can be attributed to anthropometrics and maturation changes or academic commitments besides other factors. However, it is not surprising that no swimmer could move straight from Cluster 2 to Cluster 1. The change of a swimmer from a high stability cluster to a relatively low stability cluster implies that the swimmer could not improve his or her performance so much as the counterparts did (i.e. between-subject comparison). So, with fine-tuning of the cluster membership, the number of "talented" swimmers may dwindle. Some "talented" swimmers, at some point, who fail to stay in that cluster, drop to the "proficient" cluster and similarly those in the "proficient" cluster to the "no-proficient" cluster. With this, the typical pyramid shape of the selection process was verified in M3, and in a way in M2 also. At the base of the pyramid were the "no-proficient" swimmers (N = 18), in the middle the "proficient" swimmers (N = 9) and at the top the "talented" swimmers (N = 6). Holistic research encompassing motor control, training (i.e. sports periodization), biomechanics and physiology can, in the near future, bring more insight into this phenomenon.

## **Practical implications**

The technique presented here is an important step to identify, classify and follow-up young athletes. This technique allows to assess changes in performance over time, and how the assessment can be related to the changes in the partial contribution of the determinant factors, or at least a set of factors. We showcase this procedure with young swimmers, although it can be applied across several sports, ages and competitive levels (i.e. including adult/elite sportsmen).

Performance, particularly in competitive swimming, is a multi-dimensional phenomenon, characterized by a highly complex interplay between several variables. We were able to successfully classify young swimmers, based on anthropometrics, kinematics, hydrodynamics and efficiency. It was also possible determine if their performance depended more on intrinsic (i.e. anthropometrics and biological development) or extrinsic (i.e. technique enhancement and training) factors. However, we found that the main factors, explaining performance change over time. Hence, the main determinant factors, or set of factors, explaining the performance at any given moment might not hold good for the preceding or following moment. By adopting this procedure, coaches and sports analysts will gain also insight about the possible drop rate or at least the likelihood of changes in cluster memberships over time. This technique enables the sports practitioners to design customized training sessions for each group.

Another potential use for this method is to be able to classify those swimmers who are more likely to be responsive to training or interventions programs. Therefore one might check if they are more likely to keep progressing their performance after the growth and maturation period or not. This same procedure can be selected in a near future for talent ID.

## **Conclusion**

Cluster stability is a feasible, comprehensive and informative method to gain insight into young swimmers changes over time. They can be classified into different clusters, based on their performance and determinant factors. Overall, along the

season, it was found that the stability was moderate and that the contribution of each performance determinant factor, or set of factors, may change over time.

## References

Barbosa TM, Costa M, Marinho DA, Coelho J, Moreira M, Silva AJ. (2010a). Modeling the links between young swimmer's performance: energetic and biomechanical profiles. *Pediatr Exerc Sci.* 22: 379-391.

Barbosa TM, Bragada JA, Reis VM, Marinho DA, Carvalho C, Silva AJ. (2010b). Energetics and biomechanics as determining factors of swimming performance: updating the state of the art. *J Sci Med Sport.* 13: 262-269.

Barbosa TM, Morouço PG, Jesus S, Feitosa WG, Costa MJ, Marinho DA, Silva AJ, Garrido ND. (2013). The interaction between intra-cyclic variation of the velocity and mean swimming velocity in young competitive swimmers. *Int J Sports Med.* 34: 123-130.

Bideault G, Herault R, Seifert L. (2013). Data modeling reveals inter-individual variability of front crawl swimming. *J Sci Med Sport.* 16: 281-285.

Chen I, Chen MY, Jin C, Yan H. (2008). Large-scale cluster analysis of elite male and female swimmers' race pattern. *Int J Sports Sci Eng.* 2: 123-128.

Chow JY, Davids K, Button C, Rein R. (2008). Dynamics of movement patterning in learning a discrete multiarticular action. *Motor Control.* 12: 219-240.

Clatworthy J, Buick D, Hankins M, Weinman J, Horne R. (2005). The use and reporting of cluster analysis in health psychology: a review. *Br J Health Psychol.* 10: 329-358.

Costa MJ, Marinho DA, Bragada JA, Silva AJ, Barbosa TM. (2011). Stability of elite freestyle performance from childhood to adulthood. *J Sports Sci.* 29: 1183-1189.

Costill DL, Kowaleski J, Porter D, Kirwan R, Fielding R, King D. (1985). Energy expenditure during front crawl swimming: predicting success in middle-distance events. *Int J Sports Med.* 6: 266-270.

Craig A, Pendergast D. (1979). Relationships of stroke rate, distance per stroke and velocity in competitive swimming. *Med Sci Sports*. 11: 278-283.

Durand-Bush N, Salmela J. (2002). The development and maintenance of expert athletic performance: perceptions of World and Olympic champions. *J Appl Sport Psychol*. 14: 154-171.

Ferguson CJ. (2009). An effect size primer: a guide for clinicians and researchers. *Professional Psychology: Research and Practice*. 40: 532-538.

Handl J, Knowles J, Kell DB. (2005). Computational cluster validation in post-genomic data analysis. *Bioinformatics*. 21: 3201-3212.

Hopkins WG, Hewson DJ. (2001). Variability of competitive performance of distance runners. *Med Sci Sports Exerc*. 33: 1588-1592.

Kolmogorov S, Duplisheva O. (1992). Active drag, useful mechanical power output and hydrodynamic force in different swimming strokes at maximal velocity. *J Biomech*. 25: 311-318.

Komar J, Sanders RH, Chollet D, Seifert L. (2014). Do qualitative changes in inter-limb coordination lead to effectiveness of aquatic locomotion rather than efficiency? *J Appl Biomech*. 30(2): 189-196.

Lätt E, Jürimäe J, Haljaste K, Cicchella A, Purge P, Jürimäe T. (2009). Longitudinal development of physical and performance parameters during biological maturation of young male swimmers. *Percept Mot Skills*. 108: 297-307.

Liu Y-T, Mayer-Kress G, Newell KM. (2006). Qualitative and quantitative change in the dynamics of motor learning. *J Exp Psychol*. 32: 380-393.

Marinho DA, Barbosa TM, Costa MJ, Figueiredo C, Reis VM, Silva AJ, Marques MC. (2010). Can 8 weeks of training affect active drag in young swimmers? *J Sports Sci Med*. 9: 71-78.

Milligan GW. (1981). A review of Monte Carlo tests of cluster analysis. *Multivariate Behav Res*. 16: 379-407.

Morais JE, Costa MJ, Mejias EJ, Marinho DA, Silva AJ, Barbosa TM. (2011). Morphometric study for estimation and validation of trunk transverse surface area to assess human drag force on water. *J Hum Kinet.* 28: 5-13.

Morais JE, Jesus S, Lopes V, Garrido ND, Silva AJ, Marinho DA, Barbosa TM. (2012). Linking selected kinematic, anthropometric and hydrodynamic variables to young swimmer performance. *Pediatr Exerc Sci.* 24: 649-664.

Morais JE, Saavedra JM, Costa MJ, Silva AJ, Marinho DA, Barbosa TM. (2013). Tracking young talented swimmers: follow-up of performance and its biomechanical determinant factors. *Acta Bioeng Biomech.* 15: 129-138.

Neiva HP, Marques MC, Fernandes RJ, Viana JL, Barbosa TM, Marinho DA. (2014). Does warm-up have a beneficial effect on 100-m freestyle? *Int J Sports Physiol Perform.* 9: 145-150.

Northstone K, Smith AD, Newby PK, Emmett PM. (2012). Longitudinal comparisons of dietary patterns derived by cluster analysis in 7-to 13-year-old children. *Br J Nutr.* 15: 1-9.

Phillips E, Davids K, Renshaw I, Portus M. (2010). Expert performance in sport and the dynamics of talent development. *Sports Med.* 40: 271-283.

Rein R, Button C, Davids K, Summers J. (2010). Cluster analysis of movement pattern dynamics in multi-articular actions. *Motor Control.* 14: 211-239.

Rein R. (2012). Measurement methods to analyze changes in coordination during motor learning from a non-linear perspective. *Open Sports Sci J.* 5: 36-48.

Seifert L, Chollet D. (2009). Modelling spatial-temporal and coordinative parameters in swimming. *J Sci Med Sport.* 12: 495-499.

Seifert L, Leblanc H, Herault R, Komar J, Button C, Chollet D. (2011). Inter-individual variability in the upper-lower limb breaststroke coordination. *Hum Mov Sci.* 30: 550-565.

Seifert L, Komar J, Crettenand F, Dadashi F, Aminian K, Millet GP. (2014). Inter-limb coordination and energy cost in swimming. *J Sci Med Sport*. 17(4): 439-444.

Vantorre J, Seifert L, Fernandes RJ, Boas JP, Chollet D. (2010). Kinematical profiling of the front crawl start. *Int J Sports Med*. 31: 16-21.

Zamparo P, Pendergast DR, Mollendorf J, Termin A, Minetti AE. (2005). An energy balance of front crawl. *Eur J Appl Physiol*. 94: 134–144.

Zamparo P. (2006). Effects of age and gender on the propelling efficiency of the arm stroke. *Eur J Appl Physiol*. 97: 52-58.

---

## **Chapter 4**

### **Determinant factors of long-term performance development in young swimmers**

---



## Abstract

The aims of this study were: (i) develop a performance predictor model based on the swimmers' biomechanical profile; (ii) relate the partial contribution of the main predictors with the training program and; (iii) analyze the time effect, sex effect and time X sex interaction. Ninety one swimmers (44 boys:  $12.04 \pm 0.81$  years, 47 girls:  $11.22 \pm 0.98$  years) were evaluated during a 3 year period. The decimal age, anthropometric, kinematic and efficiency features were collected in ten different moments over three seasons (i.e. longitudinal research). Hierarchical linear modeling was the procedure used to estimate the performance predictors. Results: Performance improved between season #1 - early and season #3 - late for both sexes (boys: 26.9% [20.88;32.96]; girls: 16.1% [10.34;22.54]). The decimal age (Estimate: -2.05;  $P < 0.001$ ), arm span (Estimate: -0.59;  $P < 0.001$ ), stroke length (Estimate: 3.82;  $P = 0.002$ ) and propelling efficiency (Estimate: -0.17;  $P = 0.001$ ) entered in the final model. Our results showed that over three consecutive seasons young swimmers' performance improved. Performance is a multifactorial phenomenon where anthropometrics, kinematics and efficiency were the main determinants. The change of these factors over time was coupled with the training plans of this talent ID program.

**Key-words:** *kinematics, anthropometrics, biomechanical predictors, contribution, talent ID*



## Introduction

These days, talent identification and development (ID) is one the main topics in sports performance for both researchers and practitioners. Identifying a potential elite sportsman at an early age is challenging (Morais et al., 2015). The talent ID process in swimming should hold three main components, as in other sports: (i) identification - identifying the athletes with the potential to reach the highest performance in adulthood and the main traits related to it (Delextrat et al., 2015); (ii) development - understand the changes in the performance and determinant factors according to training program (Matthys et al., 2013); (iii) and follow-up - learn about the changes in the performance and determinant factors during a time-frame (Mara et al., 2015).

Swimming is a multifactorial sport, where interactions between several scientific factors from different fields of science do happen. Hence, talent development and follow-up depends on genetics and environmental conditions, as well as its interactions (Barbosa et al., 2010a). The former is mainly related to genetic profiling and/or anthropometric assessment (Costa et al., 2009). The later can be monitored by control tests. A well-designed training plan can build-up physiological parameters and/or enhance the technique with a positive effect on the performance (Morais et al., 2014). However, evidence on this with youth is scarce. It is claimed that several determinant factors have different partial contributions to performance (Morais et al., 2014). However, so far little insight was gathered about these partial contributions in swimming or even in any other sport. Cross-sectional studies report that, at least for young swimmers, the biomechanics and physiology may explain up to 80% of the performance (Barbosa et al., 2010b). Moreover one study reports that biomechanics alone (including anthropometrics, hydrodynamics and kinematics) explain 60% and seems to be the main determinant field (Morais et al., 2012). However, during a season, the training program (i.e. external training load) relies on different parameters, that have an effect on the swimmers' response (i.e. internal training load) (Morais et al., 2014). The performance can depend upon different anthropometric, kinematic or efficiency features over a full season. Moreover, this might be a dynamic relationship with systematic shifts in the interplay among these factors. Nevertheless, little is known about such hypothetical relationships between internal and external training loads in young athletes.

The best way to gather insight on such relationships is based on longitudinal studies, despite in competitive swimming the vast majority are cross-sectional designs. Regarding the few papers reporting changes in young swimmers over time, there are a couple of concerns (Batalha et al., 2013; Toubekis et al., 2011; Strzala and Tyka, 2007): (i) the sample (i.e. small and underpowered samples; the subjects recruited are not always talented swimmers); (ii) the time-frame (i.e. short time-frames from few weeks up to one full season); (iii) follow-up studies (i.e. do not report or establish a relationship between internal and external training load over time). Indeed it was suggested earlier that longitudinal studies in competitive swimming should adopt the best practices of other scientific fields (Costa et al., 2012). Having said that, we failed to find in the literature a longitudinal research reporting the relationships between talent development and training program in a large sample of subjects over a long period of time.

The aims of this study were to: (i) test a performance predictor model based on the swimmers' biomechanical profile, over three consecutive seasons; (ii) relate the partial contribution of the main predictors with the training program over time and; (iii) analyze the time effect, sex effect and time X sex interaction. It was hypothesized that the partial contribution of each determinant factor might be related to the training program. A time and sex effect, and a time X sex interaction should be verified.

## **Methods**

### **Subjects**

Ninety one young swimmers (44 boys:  $217.7 \pm 69.5$  FINA points at short-course meters 100-m freestyle; and 47 girls:  $277.7 \pm 68.7$  FINA points at short-course meters freestyle) racing on regular basis at regional and national competitions were evaluated during 3 full seasons (3 years). The swimmers were under a talent identification, development & follow-up scheme, including age-group national record holders, age-group national champions, besides others. At the baseline, boys had  $12.04 \pm 0.81$  years and girls  $11.22 \pm 0.98$  years, and they had  $3.18 \pm 0.62$  years of training experience. Between the first and third seasons, they had  $5.10 \pm 1.08$ ,  $5.5 \pm 1.26$  (ranging from 3 to 7 in the season),  $7.1 \pm 1.11$  (ranging from 6 to 9 in the season) weekly training sessions, respectively. Sessions included warm-up,

recovery, slow, medium and intense pace, technical drills (Maglisho, 2003), as well as dry-land strength and conditioning sessions (twice per week) according to the training program (Figure 1).

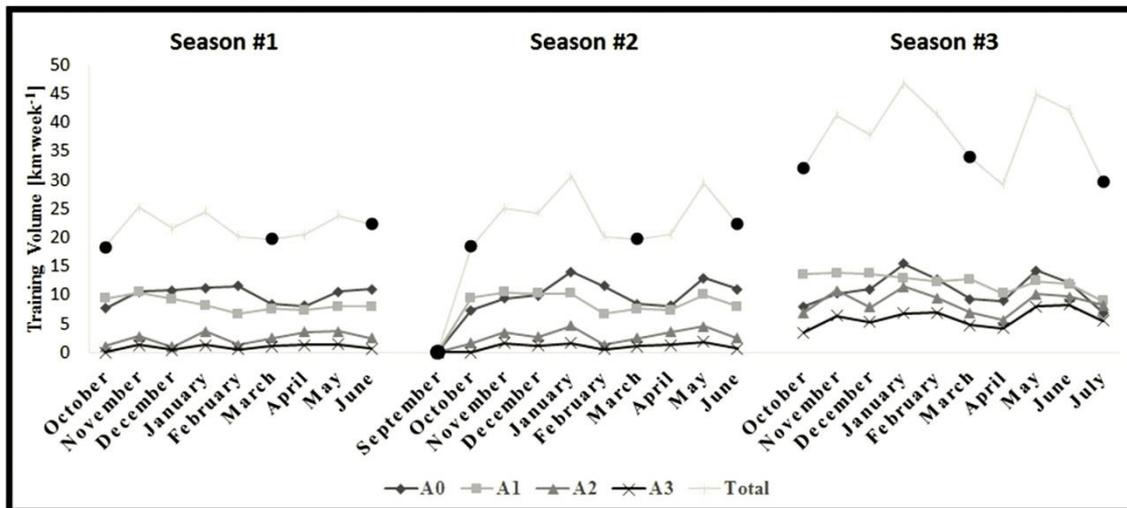


Figure 1. Training volume per week (in km) in each season, and the performance variation. ● – evaluation moments ( $M_i$ ); A0 – warm-up and recovery pace; A1 – slow pace; A2 – moderate pace (aerobic capacity); A3 – intense pace (aerobic power). For each training zone, the coefficient of variation was in season #1: 15% (A0), 14% (A1), 44% (A2), 54% (A3); season #2: 22% (A0), 16% (A1), 39% (A2), 53% (A3) and; season #3: 25% (A0), 13% (A1), 25% (A2), 26% (A3), respectively.

Coaches, parents and/or guardians and the swimmers gave the informed consent/assent to participate on this study. All procedures were in accordance to the Helsinki Declaration regarding Human research. The University of Trás-os-Montes and Alto Douro Ethic committee also approved the study design (ethic review: UTAD-2011-219).

### Study design

Repeated measures of anthropometrics, kinematics and efficiency parameters over ten different moments ( $M_i$ ), along three seasons, were performed (Figure 2). The evaluation moments were different in each season according to coaches' advices. Evaluation moments were set according to the training program and the competitive calendar in each season.

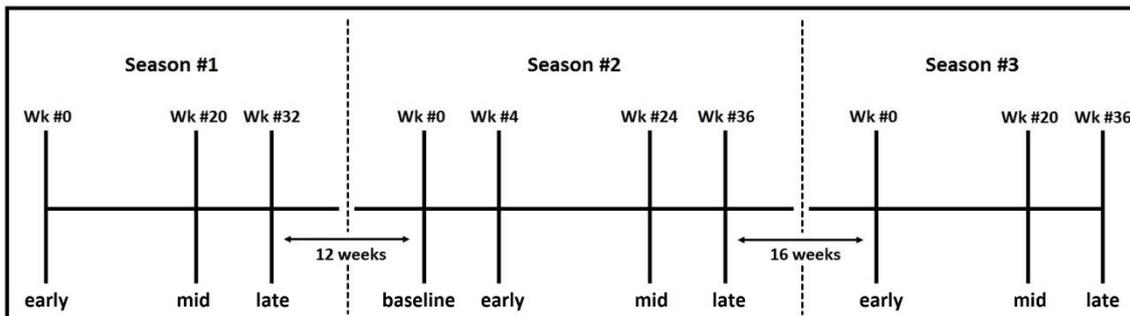


Figure 2. The timeline for the data collection over the three seasons (10 evaluation moments). All moments included the performance, kinematics, efficiency and anthropometrics assessment; #Wk – week number in each season;  $\leftrightarrow$  number of weeks break between seasons.

### Performance data collection

The 100-m freestyle event was selected as the main outcome (official race time at regional or national short course meter event). The time gap between data collection and the race was no more than two weeks.

### Kinematic data collection

The swimmers were instructed to perform three maximal freestyle swim trials of 25-m with push-off start. Between each trial, they had a 30 minutes rest to ensure a full recovery. For further analysis the average value of the three trials were calculated.

Kinematic data was collected with a mechanical technique (Swim speedo-meter, Swimsportec, Hildesheim, Germany) (ICC = 0.95). A 12-bit resolution acquisition card (USB-6008, National Instruments, Austin, Texas, USA) transferred data ( $f = 50$  Hz) to a software customized by our group (LabVIEW® interface, v.2009) (Barbosa et al., 2015). Data was exported to a signal processing software (AcqKnowledge v.3.9.0, Biopac Systems, Santa Barbara, USA) and filtered with a 5 Hz cut-off low-pass 4<sup>th</sup> order Butterworth filter. The swimming velocity ( $v$ ; in  $\text{m}\cdot\text{s}^{-1}$ ) was calculated as  $v=d/t$  in the middle 15-m (i.e. between the 5<sup>th</sup> and the 20<sup>th</sup> meter). Two experts evaluators measured the stroke frequency (SF;  $\text{cycles}\cdot\text{min}^{-1}$ ; ICC = 0.98) with a stroke counter (base 3) and then converted to SI units (Hz). The stroke length (SL; in m) was calculated as  $SL = v/SF$  (Craig and Pendergast, 1979). The intra-cyclic velocity fluctuation ( $dv$ ; dimensionless) was calculated as (Barbosa et al., 2015):

$$dv = \frac{\sqrt{\sum_i (v_i - \bar{v})^2 F_i / n}}{\sum_i v_i F_i / n} \quad (1)$$

Where  $dv$  is the intra-cyclic velocity fluctuation (dimensionless),  $v$  is the mean swimming velocity (in  $\text{m}\cdot\text{s}^{-1}$ ),  $v_i$  is the instant swimming velocity (in  $\text{m}\cdot\text{s}^{-1}$ ),  $F_i$  is the absolute frequency and  $n$  is the number of observations. The  $dv$  is a feasible way to analyze the swimmers' overall stroke mechanics, as it measures the ratio between the acceleration and deceleration within each stroke cycle, allowing to: identify critical points in the different phases of each cycle, and collect relevant data for practitioners and coaches (Barbosa et al., 2015).

### Efficiency data collection

The propelling efficiency ( $\eta_p$ ; in %) was estimated as (Zamparo, 2006):

$$\eta_p = \left[ \left( \frac{v \cdot 0.9}{2\pi \cdot SF \cdot l} \right) \cdot \frac{2}{\pi} \right] \cdot 100 \quad (2)$$

Where  $\eta_p$  is the arm's propelling efficiency (in %),  $v$  is the average swimming velocity of the swimmer (multiplied by 0.9 to take into account that, in the front crawl, about 10% of forward propulsion is produced by the legs) (in  $\text{m}\cdot\text{s}^{-1}$ ),  $SF$  is the stroke frequency (in Hz) and the term  $l$  is the average shoulder-to-hand distance (in m, i.e. this distance was measured on dry-land, while the swimmer was simulating a stroke cycle: (i) between the acromion and the olecranon; (ii) and between the olecranon and the tip of the 3<sup>rd</sup> finger, with a measuring tape (RossCraft, Canada); ICC = 0.99). The stroke index (SI; in  $\text{m}^2\cdot\text{s}^{-1}$ ) was calculated as  $SI = v \cdot SL$  (Costill et al., 1985).

### Anthropometrics data collection

All measurements were carried-out in a regular textile swimsuit, wearing cap and goggles. The body mass (BM) was measured with the swimmers in the upright position with a digital weighting scale (SECA, 884, Hamburg, Germany). The height (H) was measured in the anthropometrical position from vertex to the floor with a digital stadiometer (SECA, 242, Hamburg, Germany). The arm span (AS) was measured with swimmers standing in the upright position, arms and fingers fully extended in lateral abduction at a 90° angle with the trunk. The distance between the third fingertip of each hand was measured with a flexible anthropometric measuring tape (RossCraft, Canada) (ICC = 0.99).

## Statistical analysis

The linearity, normality and homoscedasticity assumptions were checked beforehand. Descriptive statistics included the mean, one standard deviation and the difference between first and last evaluation moment (delta), and 95% confidence interval. For the assessment of the mean stability, after running ANOVA repeated measures, Bonferroni test ( $P \leq 0.05$ ) was used to test the pairwise between the first and last evaluation moment (Costa et al., 2011). Normative stability was analyzed with Pearson's auto-correlation coefficient ( $P < 0.05$ ). As rule of thumb, for qualitative assessment, it was set that the stability was: (i) high if  $r \geq 0.60$ ; (ii) moderate if  $0.30 \leq r < 0.60$  and; (iii) low if  $r < 0.30$  (Costa et al., 2011). The longitudinal data analysis was performed by the hierarchical linear modeling (HLM). Two models were computed. The first model included the time effect, the sex effect and the time X sex interaction, to understand if: (i) there were any changes over time; (ii) differences between sexes and; (iii) differences in the changes between sexes, respectively. In the second model, decimal age, anthropometrics, kinematics and efficiency variables were tested as potential predictors. The final model only included significant predictors. Maximum likelihood estimation was calculated with the HLM5 software (Raudenbush et al., 2001).

## Results

Overall all variables showed an improvement between the first evaluation moment (season #1 - early) and the last moment (season #3 – late) (Table 1 and 2). Both boys ( $\Delta = 26.9\%$ , 95CI: [20.88;32.96],  $P < 0.001$ ) and girls ( $\Delta = 16.1\%$ , 95CI: [10.34;22.54],  $P = 0.002$ ) enhanced their performance (Table 2). Both sexes increased their BM and H. The BM was the variable with the highest difference between season #1 – early and season #3 – late (boys: 21.1%, 95CI: [15.24;26.99],  $P < 0.001$ ; girls: 16.7%, 95CI: [12.43;21.45],  $P < 0.001$ ) (Table 1). Overall, the kinematics improved in both sexes. For the boys, the v was the variable with the best improvement ( $\Delta = 17.8\%$ , 95CI: [9.00;26.60],  $P = 0.05$ ), while girls presented a meaningful, but not significant decrease in their dv ( $\Delta = -40.8\%$ , 95CI: [-69.96;-10.75],  $P = 0.64$ ), the later one suggesting a high variability (Table 2). Regarding swimming efficiency, boys and girls presented a higher improvement in the SI (boys: 24.9%, 95CI: [12.75;38.75],  $P = 0.03$ ; girls: 32.7%, [21.04;45.83],  $P = 0.001$ ). The

performance revealed a moderate-high normative stability for the boys ( $r = 0.51$ ,  $P = 0.09$  at season #1 – mid vs season #3 – mid;  $r = 0.74$ ,  $P < 0.001$  at season #2 – mid vs season #2 – late) and low-high for the girls ( $r = 0.20$ ,  $P = 0.46$  at season #1 – early vs season #3 – late;  $r = 0.95$ ,  $P < 0.001$  at season #2 – mid vs season #2 – late). As for the boys and girls pooled together, a moderate-high normative stability was observed ( $r = 0.38$ ,  $P = 0.04$  at season #1 – early vs season #3 – late;  $r = 0.98$ ,  $P < 0.001$  at season #3 – mid vs season #3 – late). Hence, wider the time-lag between evaluation moments, lower the stability is.

The HLM procedure included two stages: (1st) assess hypothetical effects/interactions in the performance with time and sex (Table 3-Model 1); (2nd) assess hypothetical relationships between changes in the performance over time with potential determinant factors (Table 3-Model 2). The results of the first hierarchical linear model tested showed that boys and girls differ significantly at the baseline (Table 3-Model 1). Girls' performance at the 100-m freestyle event was estimated as being 83.47s and boys 77.75s. The performance improved significantly over the 3 seasons (i.e. time effect). Between evaluation moments the performance improved by 1.32s. The performance enhancement was significantly higher in the boys (i.e. time X sex interaction effect). Between each moment, the performance was estimated to be higher for the boys (i.e. less 0.50s to cover the distance in comparison to girls). Therefore, time and sex have significant effects on the swimming performance.

Table 1. Descriptive statistics and variation (%; 95% CI) of the anthropometrics between season #1 - early and season #3 - late.

		Season #1			Season #2			Season #3			$\Delta$ [95% CI] Season #1 – early v season #3 - late	
		early (#1) Mean $\pm$ SD	mid (#1) Mean $\pm$ SD	late (#1) Mean $\pm$ SD	base (#2) Mean $\pm$ SD	early (#2) Mean $\pm$ SD	mid (#2) Mean $\pm$ SD	late (#2) Mean $\pm$ SD	early (#3) Mean $\pm$ SD	mid (#3) Mean $\pm$ SD		late (#3) Mean $\pm$ SD
<b>BM</b> [kg]	<b>Boys</b>	47.2 $\pm$ 10.1	48.4 $\pm$ 9.6	50.1 $\pm$ 10.0	49.7 $\pm$ 8.5	50.5 $\pm$ 8.4	52.1 $\pm$ 8.0	53.1 $\pm$ 7.6	57.9 $\pm$ 8.3	60.0 $\pm$ 7.9	59.5 $\pm$ 7.5	21.1% [15.24;26.99]
	<b>Girls</b>	44.9 $\pm$ 7.6	45.5 $\pm$ 7.8	47.2 $\pm$ 7.8	46.0 $\pm$ 7.8	46.9 $\pm$ 7.8	48.2 $\pm$ 7.9	49.0 $\pm$ 7.8	52.7 $\pm$ 6.5	53.8 $\pm$ 6.4	54.0 $\pm$ 6.6	16.7% [12.43;21.45]
<b>H</b> [cm]	<b>Boys</b>	156.9 $\pm$ 11.0	158.8 $\pm$ 10.9	159.7 $\pm$ 10.6	160.3 $\pm$ 8.5	161.6 $\pm$ 8.2	163.5 $\pm$ 8.2	164.6 $\pm$ 8.1	168.6 $\pm$ 8.2	171.0 $\pm$ 7.4	171.7 $\pm$ 7.1	8.6% [6.18;11.15]
	<b>Girls</b>	153.9 $\pm$ 8.4	155.0 $\pm$ 7.6	155.4 $\pm$ 7.8	156.2 $\pm$ 6.9	156.9 $\pm$ 6.9	157.3 $\pm$ 6.7	158.2 $\pm$ 6.6	161.2 $\pm$ 6.1	162.3 $\pm$ 5.6	163.5 $\pm$ 5.5	5.8% [4.28;7.46]
<b>AS</b> [cm]	<b>Boys</b>	161.4 $\pm$ 14.0	163.6 $\pm$ 9.2	163.8 $\pm$ 14.0	165.3 $\pm$ 12.7	165.4 $\pm$ 8.8	168.0 $\pm$ 9.0	169.4 $\pm$ 9.3	174.9 $\pm$ 9.3	176.5 $\pm$ 8.9	177.4 $\pm$ 8.4	9.0% [6.05;12.22]
	<b>Girls</b>	154.1 $\pm$ 10.0	156.2 $\pm$ 7.8	156.7 $\pm$ 8.97	157.8 $\pm$ 7.42	158.3 $\pm$ 8.3	159.4 $\pm$ 7.3	160.3 $\pm$ 7.1	164.3 $\pm$ 6.4	164.8 $\pm$ 6.6	165.7 $\pm$ 7.1	6.9% [4.97;9.05]

BM – body mass; H – height; AS – arm span; CI – confidence interval

Table 2. Descriptive statistics and variation (%; 95% CI) of the technical and performance data between season #1 - early and season #3 - late.

		Season #1			Season #2			Season #3			$\Delta$ [95% CI] Season #1 – early v season #3 - late	
		early (#1) Mean $\pm$ SD	mid (#1) Mean $\pm$ SD	late (#1) Mean $\pm$ SD	base (#2) Mean $\pm$ SD	early (#2) Mean $\pm$ SD	mid (#2) Mean $\pm$ SD	late (#2) Mean $\pm$ SD	early (#3) Mean $\pm$ SD	mid (#3) Mean $\pm$ SD		late (#3) Mean $\pm$ SD
<b>SF</b> [Hz]	Boys	0.83 $\pm$ 0.06	0.86 $\pm$ 0.07	0.88 $\pm$ 0.06	0.88 $\pm$ 0.09	0.88 $\pm$ 0.10	0.91 $\pm$ 0.09	0.90 $\pm$ 0.10	0.87 $\pm$ 0.06	0.88 $\pm$ 0.06	0.90 $\pm$ 0.08	7.6% [3.57;11.01]
	Girls	0.82 $\pm$ 0.13	0.82 $\pm$ 0.09	0.80 $\pm$ 0.07	0.82 $\pm$ 0.11	0.82 $\pm$ 0.10	0.80 $\pm$ 0.08	0.81 $\pm$ 0.08	0.78 $\pm$ 0.06	0.81 $\pm$ 0.07	0.82 $\pm$ 0.08	-0.28% [-8.20;7.63]
<b>SL</b> [m]	Boys	1.55 $\pm$ 0.31	1.10 $\pm$ 0.18	1.45 $\pm$ 0.26	1.55 $\pm$ 0.19	1.58 $\pm$ 0.20	1.60 $\pm$ 0.21	1.64 $\pm$ 0.21	1.76 $\pm$ 0.15	1.76 $\pm$ 0.14	1.75 $\pm$ 0.17	11.1% [3.04;20.23]
	Girls	1.40 $\pm$ 0.34	1.12 $\pm$ 0.27	1.38 $\pm$ 0.24	1.51 $\pm$ 0.21	1.54 $\pm$ 0.20	1.66 $\pm$ 0.17	1.66 $\pm$ 0.17	1.74 $\pm$ 0.13	1.70 $\pm$ 0.14	1.73 $\pm$ 0.15	18.7% [9.30;28.95]
<b>v</b> [m·s <sup>-1</sup> ]	Boys	1.29 $\pm$ 0.22	0.95 $\pm$ 0.14	1.28 $\pm$ 0.19	1.35 $\pm$ 0.14	1.37 $\pm$ 0.13	1.44 $\pm$ 0.14	1.47 $\pm$ 0.13	1.52 $\pm$ 0.09	1.55 $\pm$ 0.07	1.56 $\pm$ 0.08	17.8% [9.00;26.60]
	Girls	1.18 $\pm$ 0.21	0.90 $\pm$ 0.16	1.11 $\pm$ 0.19	1.23 $\pm$ 0.12	1.25 $\pm$ 0.11	1.33 $\pm$ 0.11	1.33 $\pm$ 0.10	1.35 $\pm$ 0.08	1.37 $\pm$ 0.06	1.41 $\pm$ 0.07	15.7% [7.03;24.24]
<b>dv</b> [dimensionless]	Boys	0.08 $\pm$ 0.01	0.11 $\pm$ 0.05	0.08 $\pm$ 0.01	0.09 $\pm$ 0.03	0.09 $\pm$ 0.03	0.09 $\pm$ 0.01	0.09 $\pm$ 0.01	0.09 $\pm$ 0.02	0.09 $\pm$ 0.01	0.08 $\pm$ 0.02	2.1% [-20.74;15.08]
	Girls	0.11 $\pm$ 0.05	0.10 $\pm$ 0.04	0.10 $\pm$ 0.03	0.10 $\pm$ 0.03	0.09 $\pm$ 0.03	0.08 $\pm$ 0.02	0.08 $\pm$ 0.02	0.10 $\pm$ 0.04	0.09 $\pm$ 0.02	0.08 $\pm$ 0.02	-40.8% [-69.96;-10.75]
<b>SI</b> [m <sup>2</sup> ·s <sup>-1</sup> ]	Boys	2.06 $\pm$ 0.66	1.07 $\pm$ 0.36	1.90 $\pm$ 0.61	2.11 $\pm$ 0.44	2.18 $\pm$ 0.44	2.35 $\pm$ 0.48	2.43 $\pm$ 0.46	2.68 $\pm$ 0.36	2.74 $\pm$ 0.29	2.74 $\pm$ 0.37	24.9% [12.75;38.75]
	Girls	1.63 $\pm$ 0.58	1.05 $\pm$ 0.50	1.56 $\pm$ 0.51	1.87 $\pm$ 0.38	1.93 $\pm$ 0.37	2.20 $\pm$ 0.34	2.22 $\pm$ 0.34	2.36 $\pm$ 0.2	2.33 $\pm$ 0.26	2.43 $\pm$ 0.27	32.7% [21.04;45.83]
$\eta_p$ [%]	Boys	28 $\pm$ 5	20 $\pm$ 3	26 $\pm$ 4	28 $\pm$ 3	29 $\pm$ 3	32 $\pm$ 6	30 $\pm$ 4	30 $\pm$ 2	30 $\pm$ 2	29 $\pm$ 2	2% [-7.34;11.56]
	Girls	26 $\pm$ 7	21 $\pm$ 5	26 $\pm$ 5	30 $\pm$ 4	28 $\pm$ 3	35 $\pm$ 5	32 $\pm$ 5	31 $\pm$ 3	31 $\pm$ 2	31 $\pm$ 3	15% [4.71;25.56]
<b>PERF</b> [s]	Boys	76.26 $\pm$ 7.00	71.73 $\pm$ 7.29	68.88 $\pm$ 6.66	73.48 $\pm$ 8.10	69.93 $\pm$ 7.86	67.15 $\pm$ 6.94	66.33 $\pm$ 6.36	62.00 $\pm$ 3.14	60.55 $\pm$ 3.23	60.08 $\pm$ 3.22	26.9% [20.88;32.96]
	Girls	79.06 $\pm$ 6.77	74.30 $\pm$ 4.55	72.50 $\pm$ 4.11	80.32 $\pm$ 8.60	77.66 $\pm$ 8.01	74.16 $\pm$ 6.82	73.05 $\pm$ 5.72	69.70 $\pm$ 3.98	68.54 $\pm$ 3.75	68.06 $\pm$ 4.40	16.1% [10.34;22.54]

SF – stroke frequency; SL – stroke length; v – swimming velocity; dv – intra-cyclic velocity fluctuation; SI – stroke index;  $\eta_p$  – propelling efficiency; PERF – performance

Because there were significant effects/interactions, in the second model, these predictors were retained and added to the decimal age, anthropometrics, kinematics and efficiency variables selected. The second model (i.e. final model) retained as final predictors of performance the decimal age, AS, SL and  $\eta_p$  (Table 3-Model 2). In this second stage, there were no sex and time effects or time X sex interaction. So, boys and girls could be pooled together having an overall estimation of 73.75s at the 100-m freestyle (Table 3-Model 2). The decimal age, AS and  $\eta_p$  had positive effects on the performance. By increasing one unit in the decimal age (in years), performance enhanced 2.05s. For each unit increase in AS (in cm) performance improve 0.59s. Same trend for the  $\eta_p$ , for each unit increase (in %) the performance improve 0.17s. The SL was estimated as having an inverse relationship with performance. Increasing the SL in one unit (in m) the performance was predicted as decreasing by 3.82s (i.e. more time to cover the distance) (Table 3-Model 2). Hence, age, anthropometrics variables, kinematics and swim efficiency are determinant factors to enhance the performance over 3 seasons.

Table 3. Parameters of the two models computed with standard errors (SE) and 95% confidence intervals (CI).

Parameter Fixed Effect	Estimate (SE)	95% CI	P value
<b>Model 1</b>			
Intercept	83.47(1.62)	86.67 – 80.28	<0.001
Time	-1.32(0.16)	-1.00 – -1.64	<0.001
Sex	-5.72(2.23)	-1.34 – -10.10	0.01
Time X Sex	-0.50(0.23)	-0.03 – -0.97	0.035
<b>Model 2</b>			
Intercept	73.65(0.85)	75.33 – 71.97	<0.001
Decimal Age	-2.05(0.32)	-1.42 – -2.68	<0.001
AS	-0.59(0.04)	-0.50 – -0.68	<0.001
SL	3.82(1.22)	6.23 – 1.42	0.002
$\eta_p$	-0.17(0.05)	-0.06 – -0.27	0.001

Model 1 – first model computed, including only the time effect, sex effect and time X sex interaction; Model 2 – final model, retaining the final performance predictors; AS – arm span; SL – stroke length;  $\eta_p$  – propelling efficiency.

## Discussion

The aims of this study were to test a model to predict swimming performance over three seasons in young swimmers and to learn about the partial contribution of each predictors. Main finding was that performance is related to the age (decimal age), anthropometrics (AS), kinematics (SL) and efficiency ( $\eta_p$ ).

Performance improved over the 3 seasons (3 years), and the main determinants presented an overall increase. Previous studies tracking young swimmers' performance and its determinant factors reported an increase over three evaluation moments (Latt et al., 2009a,b). In the present study, the performance showed the same trend, with an overall moderate-high stability. However, if one includes more intermediate evaluations (as this study), some of the determinant factors (kinematic and efficiency) may present slight and circumstantial increases and decreases between evaluation moments (Table 2). Overall, these changes are not significant, being the model linear. This variance seems to be coupled with the training program (Figure 1). For instance, as reported earlier for one single season, it seems that for three consecutive seasons building-up aerobic capacity and technique improvement also has an effect on the kinematics and efficiency and hence on the performance (Morais et al., 2014).

Over the three years, there is an increase in the total volume and an improvement in the performance (Figure 1). Doing the breakdown of the volume into energetic bands, it is also obvious such increase in the external training load. At the beginning of each season (between the first and intermediate moments) the training program is based on high training volumes (mainly A0: warm-up and recovery pace; and A1: slow pace). It is when there is the highest improvement in the performance (season one: 6.41%; season two: 4.71%; season three: 1.68%). In the middle of each season (between the intermediate and last moments), there is an increase in the training volume at higher regimes such as aerobic capacity and power (A2 and A3, respectively). Swimmers improved their performances by 2.48% (season #1), 1.51% (season #2) and 0.70% (season #3) in such period of time. Some of these energetic regimes are coupled with the enhancement of the technique. Coaches tend to spend a lot of time with technical drills and delivering cues on the swimmers' technique, having as well a positive effect on the performance (Barroso et al., 2015; Morais et al., 2014). Therefore, it seems that there is a clear relationship between the training program designed, the external training load and the performance enhancement within each season and over consecutive seasons.

The final hierarchical model included the decimal age, AS, SL and the  $\eta_p$ . The swimmers were evaluated in a three-year period. As the swimmer gets old, happens a shift in biological maturation (season #1 and #2: Tanner 1-2; season #3: Tanner 2-3). Because we did not measure the biological maturation, the decimal age was chosen as a surrogate variable. The increase in one unit in the decimal age (in years) was related to a 2.05s improvement in the performance. The age and anthropometrics seem to be major determinants. However these are intrinsic factors that one practitioner hardly can change but should be aware and acknowledge. The SL and  $\eta_p$  also included in the model are not genetically predicted, so coaches can play a role helping swimmers to improve it. Silva and co-workers compared the kinematics and efficiency between pre-pubertal and post-pubertal swimmers with similar training background. Main findings were that post-pubertal swimmers had significantly higher  $v$ , SL and SI than younger counterparts (Silva et al., 2013).

Anthropometric features are highly associated with young swimmers' performance (Morais et al., 2015; Latt et al., 2009a,b). The AS presents a high contribution to

performance (Silva et al., 2013; Morais et al., 2012). A higher AS leads to a higher  $v$  and hence to a better performance. During the three-year assessment, one unit increment (in cm) in the AS imposed a 0.59s improvement in the performance. Surprisingly, the SL increase over time but had a negative impact on the performance. Estimations showed that for the swimmers assessed, an increase in the SL impaired the performance. Literature reports that a higher SL provides better performances, and some of that is due to a higher AS ( $r = 0.55$ ;  $P < 0.05$ ) (Morais et al., 2012), ( $r = 0.91$ ;  $P < 0.01$ ) (Saavedra et al., 2010). However, these studies are cross-sectional designs or evaluate the swimmers during a shorter time-frame. Added to that, the swimmers were not evaluated during the transition from a pre-pubertal to post-pubertal maturational stages when significant motor control changes do happen (Barnett et al., 2008). During childhood, swimmers as any other children suffer changes in kinematics and motor control patterns. Motor learning is a process of acquiring movement patterns, which satisfy the key constraints on each individual (Davids et al., 2008; Maglisho, 2003). So it seems that during the maturation stage, the swimmers “relearn” some technical features associated to motor control aspects. Wilson and Hyde pointed out an age-related variation on kinematic measures, suggesting a continual refinement of these parameters between older childhood and early adulthood (Wilson and Hyde, 2013). In opposition to the conventional demonstration, the constraint-led approach provides a framework, combining a balanced interaction between individual, environmental and task constraints (Chow et al., 2014; Davids et al., 2008). In teaching and/or swimming training, the coaches should put the focus on individual task goals instead of relying on a standard coordination pattern (Seifert et al., 2014). The need to explore different strategies to reach a given outcome in motor control lead eventually to the non-linear pedagogy framework (Chow et al., 2014; Davids et al., 2008). The later one suggests that there is more than one way to reach the same goal. Indeed, Strzala and Tyka suggested that a SL decrease may occur, and that the swimming performance enhances throughout a SF increase (Strzala and Tyka, 2007). However, in our study, the SL showed a high coefficient of variation in comparison to the remaining predictors and can be explained under the constraint-led framework as reported earlier. It can be speculated that this higher variability concurrent with the maximum likelihood estimation explains the final outcome in the model. The performance enhancement is

a multi-factorial phenomenon and relies on different features throughout a time-frame (Morais et al., 2014) and not only on the SL. Besides that, there is a significant and inverse relationship between SL and SF (Barden et al., 2011), suggesting therefore that the increase of the later parameter took place to increase the speed and ultimately to excel. Albeit these considerations, from season #3 – early onwards, the SL improved and became more stable. One might consider that probably those adjustments were acquired. However with only two measurements remains to be complete clear such trend. As for the  $\eta_p$ , one unit increase (in %) lead to a 0.17s improvement in the performance. In the training programs, a higher attention should be given to the efficiency and not only to training volume and intensity.

### **Practical Implications**

The HLM is a comprehensive and straightforward way to model young swimmers' performance. Swimming performance does not depend on isolated features but from the interaction among several (Barbosa et al., 2010a). Based on the final model, intrinsic factors, more related to “nature” (such as the decimal age and anthropometrics, in this case, arm span) and extrinsic ones linked to “nurture” (including stroke length and propelling efficiency) are determinant to excel in such early ages in swimming. Besides that, there is evidence that the changes of the determinant factors over time happen in a non-linear fashion way (there are slight improvements and impairments along the way). Talent ID programs should rely on identifying the performance determinant features in several moments of the season, how these change over time and interact. Hence, evidence-based information, about the partial contribution of each determinant factor, should be provided to coaches on regular basis (within and between seasons).

So far, to the best of our understanding no study provided a deep insight on the relationship between the development of these determinants and the training program. However, some might consider that the training level and other environmental factors (nurture) are ignored in detriment of a natural growth and maturation processes (genetics) (Brutsaert and Parra, 2009). Our data shows that the training program also has a meaningful influence on the performance and its main extrinsic determinants. The same procedure and reasoning can be applied to other sports so that one can gather insight over time of performance' main

determinants in young talented athletes under different talent ID schemes of different sports.

It can be addressed as main limitations: (i) the decimal age is a surrogate variable of sexual maturation. Lately there are increasing ethic concerns regarding the direct assessment of sexual maturation by Tanner stages due to some misconduct between practitioners and athletes. Despite that, the low variability in the maturation by the self-report and undisclosed identify as we carried out suggests that there is no effect at least for this time-frame of 3 years; (ii) the kinematics and efficiency variables were collected over 25-m trials and not the 100-m freestyle race. One might consider that to ensure a more real evolution of the kinematic and efficiency features with the performance, these parameters should have been assessed during the official race or a simulated event. However, kinematics and efficiency measured during the 25-m trial, showed an overall high-very high correlation with the 100-m performance in pilot studies. E.g. for the data collected in this research the correlation between the 25-m and 100-m performance was  $r = 0.71$  ( $P < 0.001$ ). This enabled us to select straightforward, less time-consuming and insightful procedures (e.g. mechanical speedo-meter rather than motion-capture systems) that are feasible to carry out in such a large sample size over three consecutive years.

## **Conclusion**

As a conclusion, over three consecutive seasons the performance and its determinant factors improved. Young swimmers' performance is a multifactorial phenomenon where different factors play meaningful roles. Anthropometric, kinematic and efficiency features entered in the final model as main predictors. The change of these factors over time was coupled with the training program. Therefore, talent ID programs should rely not only on the identification but also on the development of the main predictors according to a well-designed training program plan in a long-term basis.

## **References**

Barbosa TM, Bragada JA, Reis VM, Marinho DA, Silva AJ. (2010a). Energetics and biomechanics as determining factors of swimming performance: updating the state of the art. *J Sci Med Sport*. 13(2): 262-269.

Barbosa TM, Costa MJ, Marinho DA, Coelho J, Moreira M, Silva AJ. (2010b). Modeling the links between young swimmers' performance: energetic and biomechanic profiles. *Pediatr Exerc Sci*. 22(3): 379-391.

Barbosa TM, Morais JE, Marques MC, Silva AJ, Marinho DA, Kee YH. (2015). Hydrodynamic profile of young swimmers: changes over a competitive season. *Scan J Med Sci Sports*. 25: 184-196.

Barden JM, Kell RT, Kobsar D. (2011). The effect of critical speed and exercise intensity on stroke phase duration and bilateral asymmetry in 200-m front crawl swimming. *J Sports Sci*. 29(5): 517-526.

Barnett LM, Van Beurden E, Morgan PJ, Brooks LO, Beard JR. (2008). Does childhood motor skill proficiency predict adolescent fitness? *Med Sci Sports Exerc*. 40(12): 2137-2144.

Barroso R, Salgueiro DF, do Carmo EC, Nakamura FY. (2015). Training volume and repetition distance affect session rating of perceived exertion and internal load in swimmers. *Int J Sports Physiol Perform*. 10(7): 848-852.

Batalha NM, Raimundo AM, Tomas-Carus P, Barbosa TM, Silva AJ. (2013). Shoulder rotator cuff balance, strength, and endurance in young swimmers during a competitive season. *J Strength Cond Res*. 27(9): 2562-2568.

Brutsaert TD, Parra EJ. (2009). Nature versus nurture in determining athletic ability. *Med Sport Sci*. 54: 11-27.

Chow JY, Koh M, Davids K, Button C, Rein R. (2014). Effects of different instructional constraints on task performance and emergence of coordination in children. *Eur J Sports Sci*. 14(3): 224-232.

Costa AM, Silva AJ, Garrido ND, Louro H, de Oliveira RJ, Breitenfeld L. (2009). Association between ACE D allele and elite short distance swimming. *Eur J Appl Physiol*. 106(6): 785-790.

Costa MJ, Marinho DA, Bragada JA, Silva AJ, Barbosa TM. (2011). Stability of elite Freestyle performance from childhood to adulthood. *J Sports Sci*. 29(11): 1183-1189.

Costa MJ, Bragada JA, Marinho DA, Silva AJ, Barbosa TM. (2012). Longitudinal interventions in elite swimming: a systematic review based on energetics, biomechanics, and performance. *J Strength Cond Res.* 26(7): 2006-2016.

Costill DL, Kowaleski J, Porter D, Kirwan R, Fielding R, King D. (1985). Energy expenditure during front crawl swimming: predicting success in middle-distance events. *Int J Sports Med.* 6: 266–270.

Craig A, Pendergast D. (1979). Relationships of stroke rate, distance per stroke and velocity in competitive swimming. *Med Sci Sports.* 11: 278-283.

Davids K, Button C, Bennett SJ. (2008). Dynamics of skill acquisition: a constraint-led approach. Champaign, IL: Human Kinetics.

Delextrat A, Grosgeorge B, Bieuzen F. (2015). Determinants of performance in a new test of planned agility for young elite basketball players. *Int J Sports Physiol Perform.* 10(2): 160-165.

Latt E, Jürimae J, Haljaste K, Cicchella A, Purge P, Jurimae T. (2009a). Physical development and swimming performance during biological maturation in young female swimmers. *Coll Antropol.* 33:117–122.

Latt E, Jurimae J, Haljaste K, Cicchella A, Purge P, Jurimae T. (2009b). Longitudinal development of physical and performance parameters during biological maturation of young male swimmers. *Percept Mot Skills.* 108: 297–307.

Maglisho E. (2003). *Swimming fastest.* Champaign, IL: Human Kinetics.

Mara JK, Thompson KG, Pumpa KL, Ball NB. (2015). Periodization and physical performance in elite female soccer players. *Int J Sports Physiol Perform.* 10(5): 664-669.

Matthys SP, Vaeyens R, Franssen J, Deprez D, Pion J, Vandendriessche J, Vandorpe B, Lenoir M, Philippaerts R. (2013). A longitudinal study of multidimensional performance characteristics related to physical capacities in youth handball. *J Sports Sci.* 31(3): 325-334.

Morais JE, Jesus S, Lopes V, Garrido N, Silva AJ, Marinho DA, Barbosa TM. (2012). Linking selected kinematic, anthropometric and hydrodynamic variables to young swimmer performance. *Pediatr Exerc Sci.* 24(4): 649-664.

Morais JE, Marques MC, Marinho Da, Silva AJ, Barbosa TM. (2014). Longitudinal modeling in sports: young swimmers' performance and biomechanics profile. *Hum Mov Sci.* 37: 111-122.

Morais JE, Silva AJ, Marinho DA, Seifert L, Barbosa TM. (2015). Cluster stability as a new method to assess changes in performance and its determinant factors over a season in young swimmers. *Int J Sports Physiol Perform.* 10(2): 261-168.

Raudenbush S, Bryk A, Cheong YF, Congdon R. (2001). *HLM 5: hierarchical linear and nonlinear modeling.* Lincolnwood, IL: Scientific Software International Inc.

Saavedra JM, Escalante Y, Ferran AR. (2010). A multivariate analysis of performance in young swimmers. *Pediatr Exerc Sci.* 22: 135-151.

Seifert L, Komar J, Barbosa TM, Toussaint H, Millet G, Davids K. (2014). Coordination pattern variability provides functional adaptations to constraints in swimming performance. *Sports Med.* 44: 1333-1345.

Silva AF, Figueiredo P, Seifert L, Soares S, Vilas-Boas JP, Fernandes RJ. (2013). Backstroke technical characterization of 11-13 year old swimmers. *J Sports Sci Med.* 12(4): 623-629.

Strzala M, Tyka A. (2007). Shaping of physical endurance and front crawl swimming technique indices in swimmers after half-year training period. *Med Sportiva.* 11(4): 88-96.

Toubekis AG, Vasilaki A, Douda H, Gourgoulis V, Tokmakidis S. (2011). Physiological responses during interval training at relative to critical velocity intensity in young swimmers. *J Sci Med Sport.* 14(4): 363-368.

Wilson PH, Hyde C. (2013). The development of rapid online control in children aged 6-12 years: reaching performance. *Hum Mov Sci.* 32(5): 1138-1150.

Zamparo P. (2006). Effects of age and gender on the propelling efficiency of the arm stroke. *Eur J Appl Physiol.* 97(1): 52-58.

## General Discussion and Conclusions

The aim of this thesis was to identify and follow-up young swimmers' performance and its determinant factors over time. The main conclusions were: (i) young swimmers' performance depends from a set of anthropometric, kinematic, efficiency and hydrodynamic factors (chapter 1); (ii) young swimmers' performance significantly improved, and with a significant inter-variability. Different determinant factors were responsible for such improvement in each evaluation moment (chapter 2); (iii) within an age-group of prepubescent swimmers, three sub-groups with similar biomechanical characteristics were found (chapter 3), and; (iv) the predictive model, based on three consecutive seasons, included anthropometrics, kinematics and efficiency determinants, showing the multifactorial phenomenon that swimming is (chapter 4).

The research starts having as aim to identify which are the young swimmers' performance main determinants and how they interact (chapter 1). A cross-sectional structural equation model was computed based on the evidence reported in the literature (i.e. exploratory researches) (e.g. Saavedra et al., 2010; Jurimae et al., 2007; Geladas et al., 2005; Toussaint et al., 1990). Main finding was that anthropometrics, kinematics, efficiency and hydrodynamics may explain up to 62% (girls), 58% (boys), and 50% (overall), of the swimming performance at this early ages. Indeed, literature reports that some variables belonging to such branches of the biomechanics, are strongly related to young swimmers' performance (Vitor and Bohme, 2012; Barbosa et al., 2010b; Marinho et al., 2010). As far as we can understand based on this cross-sectional design, the arm span, stroke length, active drag and stroke index, are variables that one should monitor over time because of their high relationship with the performance.

The next step was to model the performance and the determinant factors reported earlier (chapter 1) inputting data collected over a full season (chapter 2). The swimming performance enhanced constantly and significantly between each one of the four evaluation moment over the entire season. However the determinant factors selected did not present the same trend, albeit literature reports a constant increase/enhancement of the performance and its determinants (Latt et al., 2009a,b). Data in chapter 2, showed that this is not the case for the determinant factors. A

given determinant factor may, in some specific moment, decrease without compromising the overall performance enhancement because other(s) may play a major role at that stage of the season. This phenomenon is based on the reasoning that swimming performance is a multifactorial phenomenon, being the performance dependent upon the interaction by several determinant factors and hence, there is more than one way to reach a given outcome (Barbosa et al., 2010a). If by any reason, a specific determinant factor impairs, the swimmer is able to maintain and/or improve the performance by the increase/enhancement of other(s) key-determinant(s).

Interestingly we did observe that in each evaluation moment, a relationship could be found between the determinant factor with the highest and significant effect to the performance and the training program designed and implemented (i.e. relationship between internal and external training load). In the first half of the season the training program was characterized by high training volumes to build-up energetics. And in the second half the swimming efficiency and stroke technique were the major goals. Indeed, the latent growth curve model showed that the training program designed accomplished such objectives as this was reflected in the model's outputs.

With this latent growth curve modeling we were also able to assess the inter-subject variability. Data in chapter 2, showed that gender had a significant effect on the performance growth, having the boys delivered better performances in comparison to the girls. Besides that, variance analysis showed significant differences between swimmers at the baseline and during the performance growth. The intercept variance was significant for all models computed, suggesting an inter-subject variability at the baseline. Not only boys differed from girls (gender effect), but also only boys, and only girls were different within each gender group. This suggests that all swimmers were different among them at baseline. The slope variance was also significant for all models, suggesting a heterogeneous growth rate of the performance and hence an inter-subject variability.

The inter-subject variability reported earlier in chapter 2, led us to hypothesize that in a group of young swimmers, sub-groups with similar characteristics might be found (chapter 3). Therefore, the main aim of chapter 3 was to identify, classify, and follow-up young swimmers, based on their performance and its determinant factors over a

competitive season, analyzing simultaneously the swimmers' stability during that time-frame. Our hypothesis was correct and accepted. Three sub-groups (i.e. cluster groups) were found and characterized as: (i) "talented"; (ii) "proficient" and; (iii) "no-proficient". Similar reasoning had been reported earlier but for cross-sectional datasets (Barbosa et al., 2014). Overall, the three sub-groups did increase/enhance all parameters assessed, including the performance. The "talented" swimmers (the best performers) were characterized by higher body dimensions and parameters related to the stroke mechanics. The "proficient" swimmers (average performances) showed reasonable stroke mechanics outputs, but a higher drag though. The "no-proficient" swimmers (poorer performances) featured a higher intra-cyclic velocity fluctuation.

The assessment of the stability allowed to determinate if the swimmers were able to shift from cluster group or not (i.e. move to a better cluster or drop to a poorer one). The "no-proficient" present the highest stability, followed by the "talented" and the "proficient". The majority of swimmers gathered in the poorer performing cluster ("no-proficient") were unable to move-up to a better cluster. Half of the "talented" swimmers were able to stay in the cluster, and the other half dropped to poorer clusters. The "proficient" swimmers were the ones with the highest number of changes. Data suggests that despite not being impossible, it is very challenging for a poor swimmer to climb the performance ladder all the way up and become a "talented" swimmer. The drop-out ratio from "talented" to non-talented is almost half of the subjects' pool per season. However, the "proficient" swimmers have the chance to either go up or down. So, if hypothetically one "talented" swimmer falls to "proficient" this should be flagged. But one should keep monitoring him because there is always a good chance of the swimmer eventually shifting to a new cluster (hopefully moving up again to the likes of other "talented" counterparts).

One concern for the evidence reported in the previous studies is the time-frame being only a full season. Yet, there is a lack of understanding of such changes over a longer period of time. So, it was decided to model the performance and its determinant factors over a long period of time (dataset for three consecutive seasons, back-to-back) (chapter 4). Endogenous (performance) and exogenous variables (determinant factors) were modeled with hierarchical linear modeling

featuring two levels (level 1: the time effect, sex effect and time X sex interaction; level 2: effect of the decimal age, anthropometric, kinematic and efficiency). Performance is related to the decimal age, arm span, stroke length and propelling efficiency. Contrarily to what happened in chapter 2 study (one season), this long-term evaluation allowed understanding that in a broader-scope the performance between boys and girls is not significantly different. This suggests that growth and maturation spurs do happen in different moments for boys and girls. But in a larger scope, when this differences do happen is not so clear yet. The final predictive model confirmed that, as for adult/elite counterparts, young swimmers' performance is also a multifactorial phenomenon. Anthropometric, kinematics and efficiency are the main determinant factors.

Data also confirmed that determinant factors may present a non-linear profile. I.e. in some circumstantial moments given determinant factors may increase/enhance and/or decrease/impair without having a negative effect on the performance enhancement. Same phenomenon was reported earlier in chapter 2.

Data gathered in this thesis will enable coaches and analysts being aware of which are the young swimmers' performance determinants, and understand how they interplay over time. Technique enhancement, should not be neglected because it has a meaningful effect on the performance determinants flagged in this set of studies. A relationship between the external and internal training load was also found. Therefore, coaches may design their training programs so that swimmers may rely on specific determinant factors to enhance their performance in specific moments of the season. Data also shows that young swimmers should be seen as unique individuals, as reflected by the significant between-subjects variability verified. For practitioners that consider too challenging to design personalized training programs for each one of their swimmers, at least to consider gather them in similar sub-groups is an advice. Follow-up is for coaches designing specific training programs according to the "needs" of each sub-group. As for analysts, for example at sports institutes or swimming federations, they should consider to monitor the determinant factors described in this thesis and follow-up their young swimmers over time. On top of that, data reported here can be used for benchmark with their own swimmers.

It can be addressed as main limitations: (i) other domains such as strength & conditioning, and motor control could enhance the power of the models' outputs. That said, to add extra variables in the models, one would need to have even larger sample sizes; (ii) cutting-edge parameters, including fractal analysis and entropy could provide some extra insight on these findings; (iii) eventually, spin-off this research project to remaining swim stroke, besides front-crawl.

It can be considered as specific conclusions of this thesis:

- Anthropometrics, kinematics, efficiency and hydrodynamics were the main determinants of young swimmers' performance (chapter 1);
- Swimming performance improved, and young swimmers showed a significant inter-variability for the performance growth over a full season. Different factors were determinants for the performance improvement in each evaluation moment (chapter 2);
- Within an age-group of swimmers, three sub-groups were found based on similar characteristics: "talented", "proficient" and "no-proficient". Overall, the trend was for the number of swimmers on the "talented" sub-group, decrease over time (chapter 3);
- The predictive model developed over the three year assessment, included anthropometrics, efficiency and kinematics as main determinants, as well as the decimal age (chapter 4).

As main conclusion, anthropometric, kinematics, efficiency and hydrodynamics are the main determinant factors in young swimmers' performance. It seems that training (mainly technical) plays a major role on their performance enhancement. The changes in the performance and determinant factors are indeed related to the training program designed and implemented. Having said that, performance enhancement over time is due to a complex interaction among different factors, not being always the same having the key-role.



## General References

Barbosa TM, Bragada JA, Reis VM, Marinho DA, Silva AJ. (2010a). Energetics and biomechanics as determining factors of swimming performance: updating the state of the art. *J Sci Med Sport*. 13(2): 262-269.

Barbosa TM, Costa MJ, Marinho DA, Coelho J, Moreira M, Silva AJ. (2010b). Modeling the links between young swimmers' performance: energetic and biomechanic profiles. *Pediatr Exerc Sci*. 22(3): 379-391.

Barbosa TM, Morais JE, Costa MJ, Gonçalves J, Marinho DA, Silva AJ. (2014). Young swimmers' classification based on kinematics, hydrodynamics, and anthropometrics. *J Appl Biomech*. 30: 310-315.

Barbosa TM, Morais JE, Marques MC, Silva AJ, Marinho DA, Kee, YH. (2015). Hydrodynamic profile of young swimmers: changes over a competitive season. *Scan J Med Sci Sports*. 25: 184-196.

Batalha NM, Raimundo AM, Tomas-Carus P, Barbosa TM, Silva AJ. (2013). Shoulder rotator cuff balance, strength, and endurance in young swimmers during a competitive season. *J Strength Cond Res*. 27(9): 2562-2568.

Costa MJ, Bragada JA, Mejjias JE, Louro H, Marinho DA, Silva AJ, Barbosa TM. (2013). Effects of swim training on energetics and performance. *Int J Sports Med*. 34: 507-513.

Erlandson MC, Sherar LB, Mirwald RL, Mafulli N, Baxter-Jones AD. (2008). Growth and maturation of adolescent female gymnasts, swimmers, and tennis players. *Med Sci Sports Exerc*. 40(1): 34-42.

Garrido ND, Marinho DA, Barbosa TM, Costa AM, Silva AJ, Pérez-Turpin JA, Marques MC. (2010). Relationships between dry land strength, power variables and short sprint performance in young competitive swimmers. *J Sports Sci Med*. 5: 240-249.

Geladas ND, Nassis GP, Pavlicevic S. (2005). Somatic and physical traits affecting sprint swimming performance in young swimmers. *Int J Sports Med*. 26: 139-144.

Jurimae J, Haljaste K, Cicchella A, Latt E, Purge P, Leppik A, Jurimae T. (2007). Analysis of swimming performance from physical, physiological and biomechanical parameters in young swimmers. *Pediatr Exerc Sci.* 19: 70-81.

Komar J, Sanders RH, Chollet D, Seifert L. (2014). Do qualitative changes in inter-limb coordination lead to effectiveness of aquatic locomotion rather than efficiency? *J Appl Biomech.* 30(2): 189-190.

Latt E, Jurimae J, Haljaste K, Cicchella A, Purge P, Jurimae T. (2009a). Physical development and swimming performance during biological maturation in young female swimmers. *Coll Antropol.* 33: 117-122.

Latt E, Jurimae J, Haljaste K, Cicchella A, Purge P, Jurimae T. (2009b). Longitudinal development of physical and performance parameters during biological maturation of young male swimmers. *Percept Mot Skills.* 108: 297-307.

Matthys SP, Vaeyens R, Franssen J, Deprez D, Pion J, Vandendriessche J, Vandorpe B, Lenoir M, Philippaerts R. (2013). A longitudinal study of multidimensional performance characteristics related to physical capacities in youth handball. *J Sports Sci.* 31(3): 325-334.

Robertson S, Woods C, Gastlin P. (2014). Predicting higher selection in elite junior Australian rules football: the influence of physical performance and anthropometric attributes. *J Sci Med Sport.* 18(5): 601-606.

Saavedra JM, Escalante Y, Ferran AR. (2010). A multivariate analysis of performance in young swimmers. *Pediatr Exerc Sci.* 22: 135-151.

Toussaint HM, de Looze M, van Rossem B, Leijdekkers M, Dignum H. (1990). The effect of growth on drag in young swimmers. *Int J Sport Biomech.* 6: 18-28.

Vitor FM, Bohme MT. (2010). Performance of young male swimmers in the 100 meters front crawl. *Pediatr Exerc Sci.* 22: 278–287.