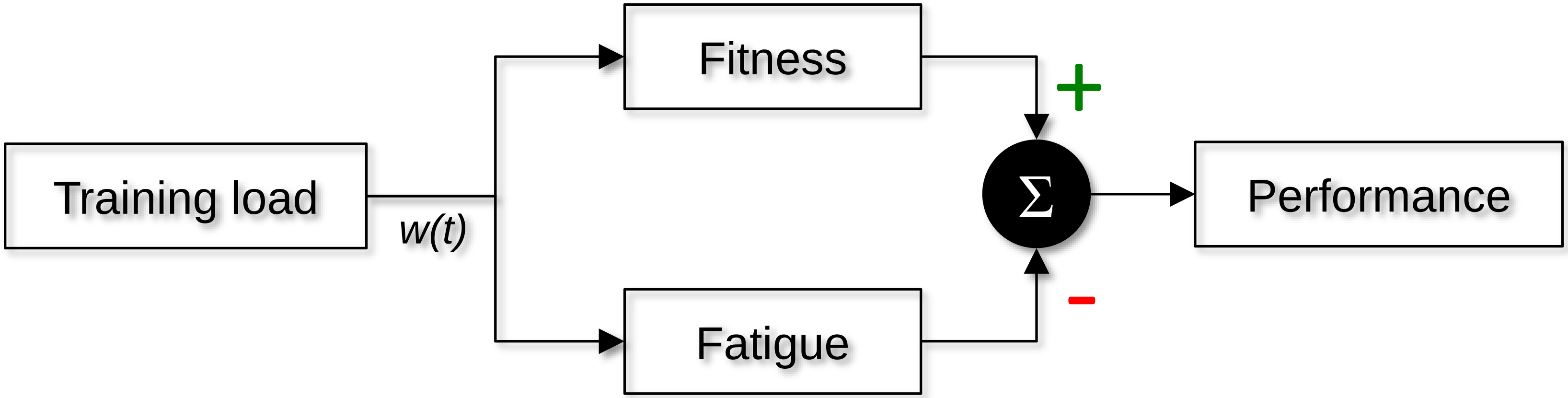
Mixed linear modelling of training-performance relationship in elite swimmers

Rodríguez F.A., Susín A., Calle M.L., Chaverri D., Iglesias X.

INEFC Institut Nacional d'Educació Física de Catalunya
UNIVERSITAT DE BARCELONA
UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH
UVIC UNIVERSITAT DE VIC UNIVERSITAT CENTRAL DE CATALUNYA

Introduction

- Banister, Calvert et al. first applied systems theory to quantifying the relationship between training and performance
- Busso et al. improved model accuracy by accounting for training-related changes on fatigue
- However, systems models have been unable to consistently predict performance on an individual basis in a “real-world” setting



Aim

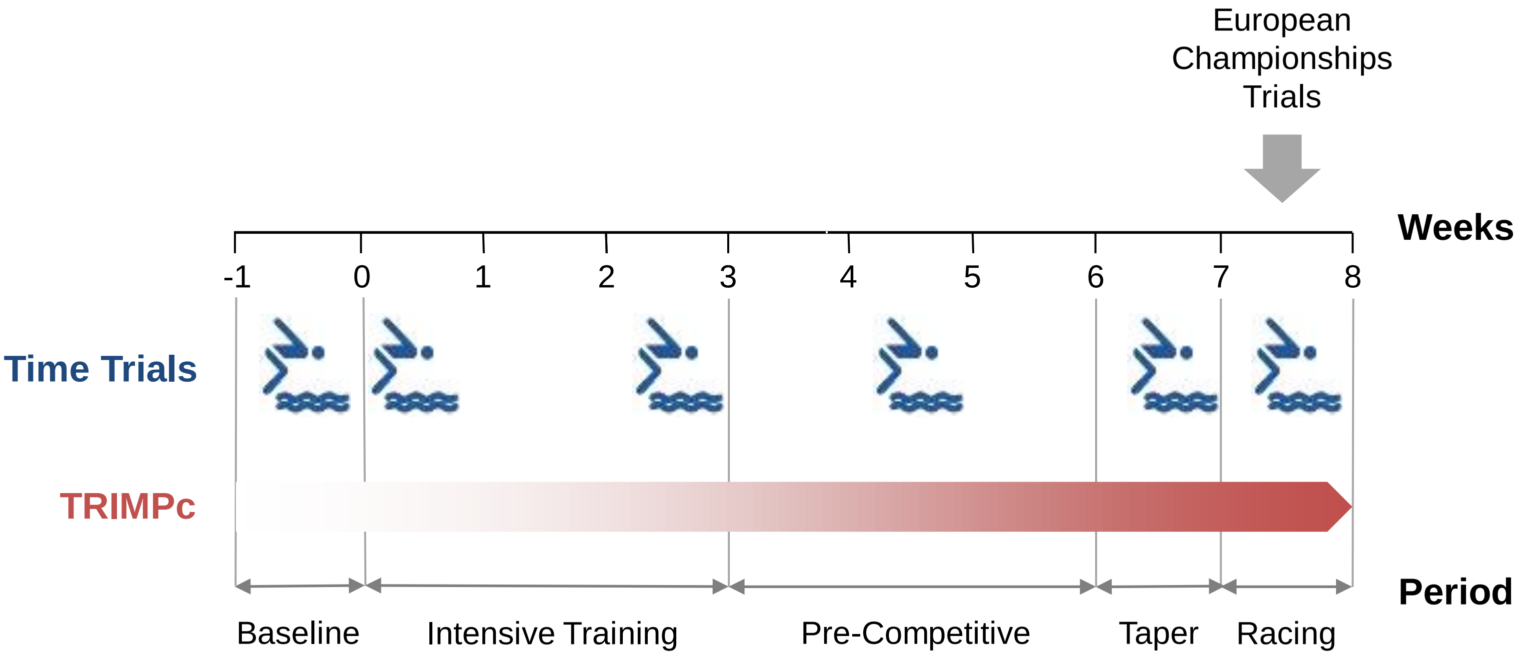
- To evaluate the modelling of the dose-response relationship between quantified training and performance with linear mixed-effects regression analysis (MixMod) in high-level competitive swimmers

- To compare the goodness-of-fit and robustness of the model with classical individual linear modeling (LinMod)

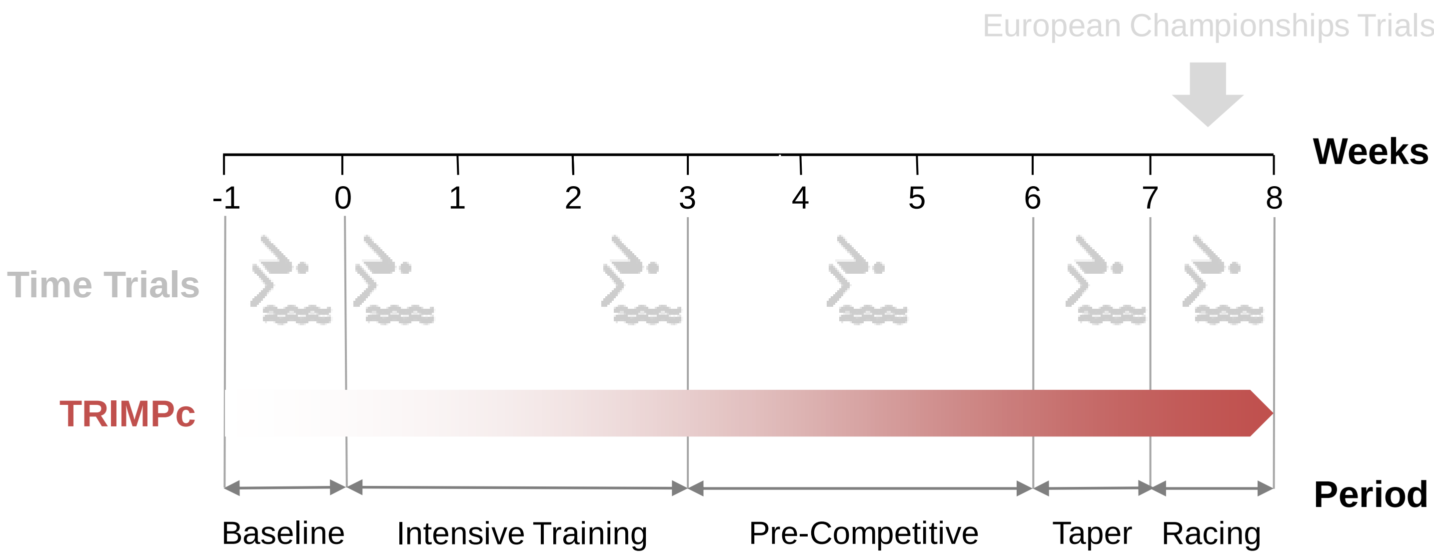
Methods

Design

- International-level swimmers (n=10)
- Training monitoring during an 8-week mesocycle



Training monitoring



- **Timing**: TX H2O (Freelap) beacon transmitters (50-m laps)
- **HR**: Water: CardioSwim monitors synced with beacons
  Dryland: Polar RS800CX monitors
- **Cummulative training impulse (TRIMPc*)**

* García-Ramos et al. (2015) Eur J Sport Sci, 15, 85-93.

Training monitoring

- **Timing**: TX H2O (Freelap) beacon transmitters (50-m laps)
- **Heart rate**:

  Water: CardioSwim monitors synced with beacons

  Dryland: Polar RS800CX monitors

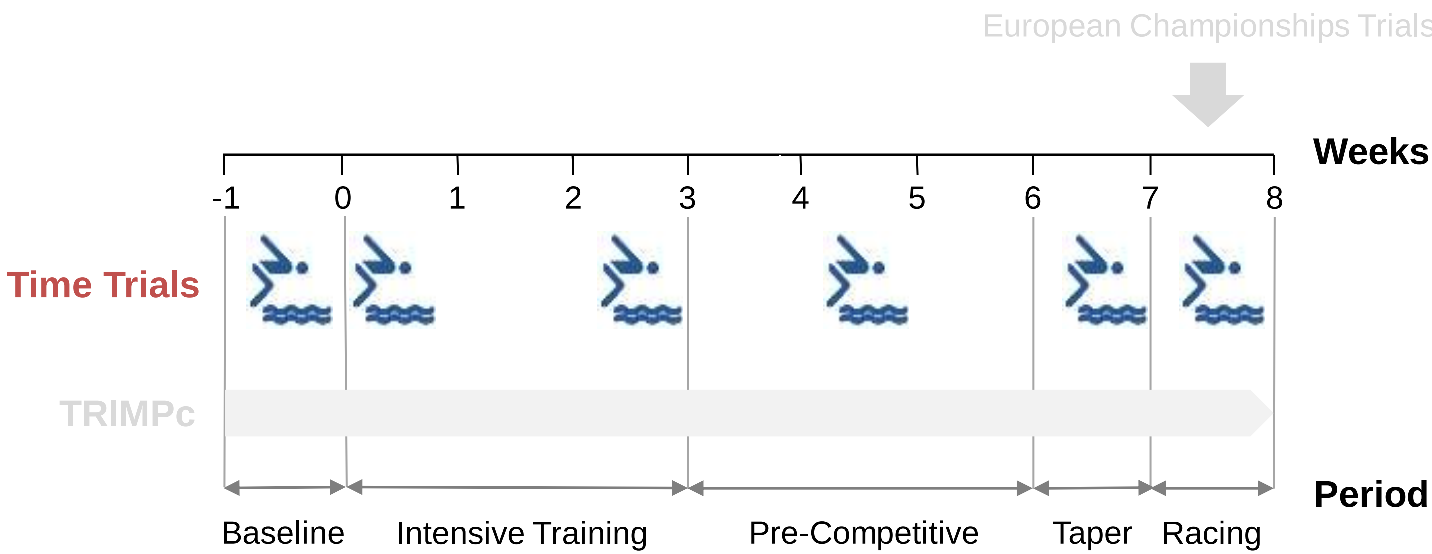
- **Cummulative training impulse (TRIMPc*)**

$$\mathrm{TRIMPc} = \sum_{i=1}^{n} t k_1 x \begin{cases} \text{for men} = k_1 = 0.64 e^{1.92x} \\ \text{for women} = k_1 = 0.86 e^{1.67x} \end{cases}$$

$$x = \Delta \mathrm{HR}_{\mathrm{ratio}} = \frac{\mathrm{HR}_{\mathrm{exercise}} - \mathrm{HR}_{\mathrm{rest}}}{\mathrm{HR}_{\mathrm{max}} - \mathrm{HR}_{\mathrm{rest}}}$$

* García-Ramos et al. (2015) Eur J Sport Sci

Performance testing



Time trials:

50 and 400 m freestyle

100 m (sprinters) or 200 m (non-sprinters) best stroke

Mathematical modelling

- **Banister-Busso* individual linear regression model (LinMod)**

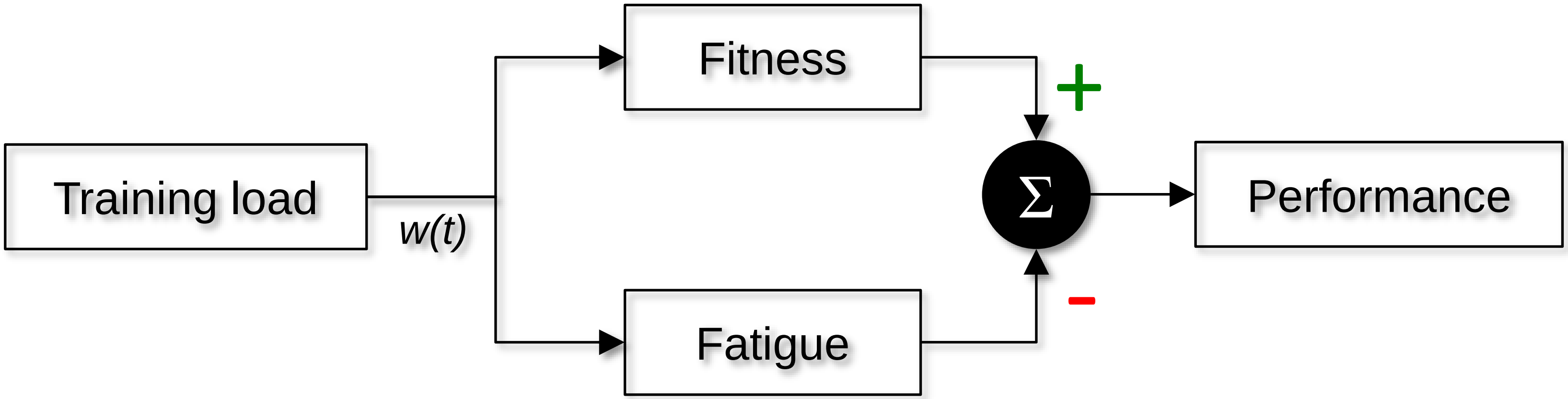
- **Linear mixed-effects regression analysis model (MixMod)**

  fixed-effects parameters: general / sample pattern

  random-effects parameters: individual specific behaviour

* Busso T (2003) Med Sci Sports Exerc, 35, 1188-1195

Banister-Busso linear model



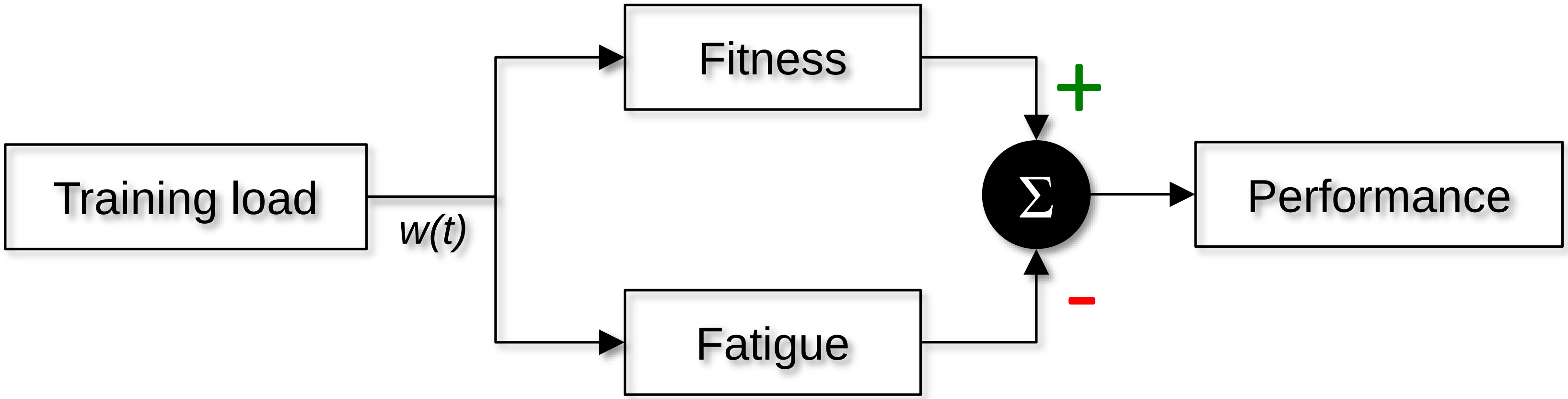
$$\hat{p}(n) = p^* + k_1 \cdot Fitness(\tau_1, w, n) + k_3 \cdot Fatigue(\tau_2, \tau_3, w, n)$$

$$\hat{p} = p^* + k_1 \sum_{i=1}^{n-1} w^i e^{-\frac{n-i}{\tau_1}} - \sum_{i=1}^{n-1} k_2^i w^i e^{-\frac{n-i}{\tau_2}}$$

$$k_2^i = k_3 \sum_{j=1}^{i} w^j e^{-\frac{i-j}{\tau_3}}$$

Banister et al (1991); Busso (2003)

Mixed linear model (MixMod)



$$y_i(t) = \beta_0 + b_{i0} + (\beta_1 + b_{i1})\cdot x_1(t) + (\beta_2 + b_{i2})\cdot x_2(t)$$

$$\hat{p} = p^* + k_1 \sum_{i=1}^{n-1} w^i e^{-\frac{n-i}{\tau_1}} - k_3 \sum_{i=1}^{n-1} \sum_{j=1}^{i} w^j e^{-\frac{n-i}{\tau_2}}$$

$\beta_0, \beta_1, \beta_2 =$ fixed (sample) effects

$b_{i0}, \beta_{i1}, \beta_{i2} =$ random (individual) effects

Statistical analysis

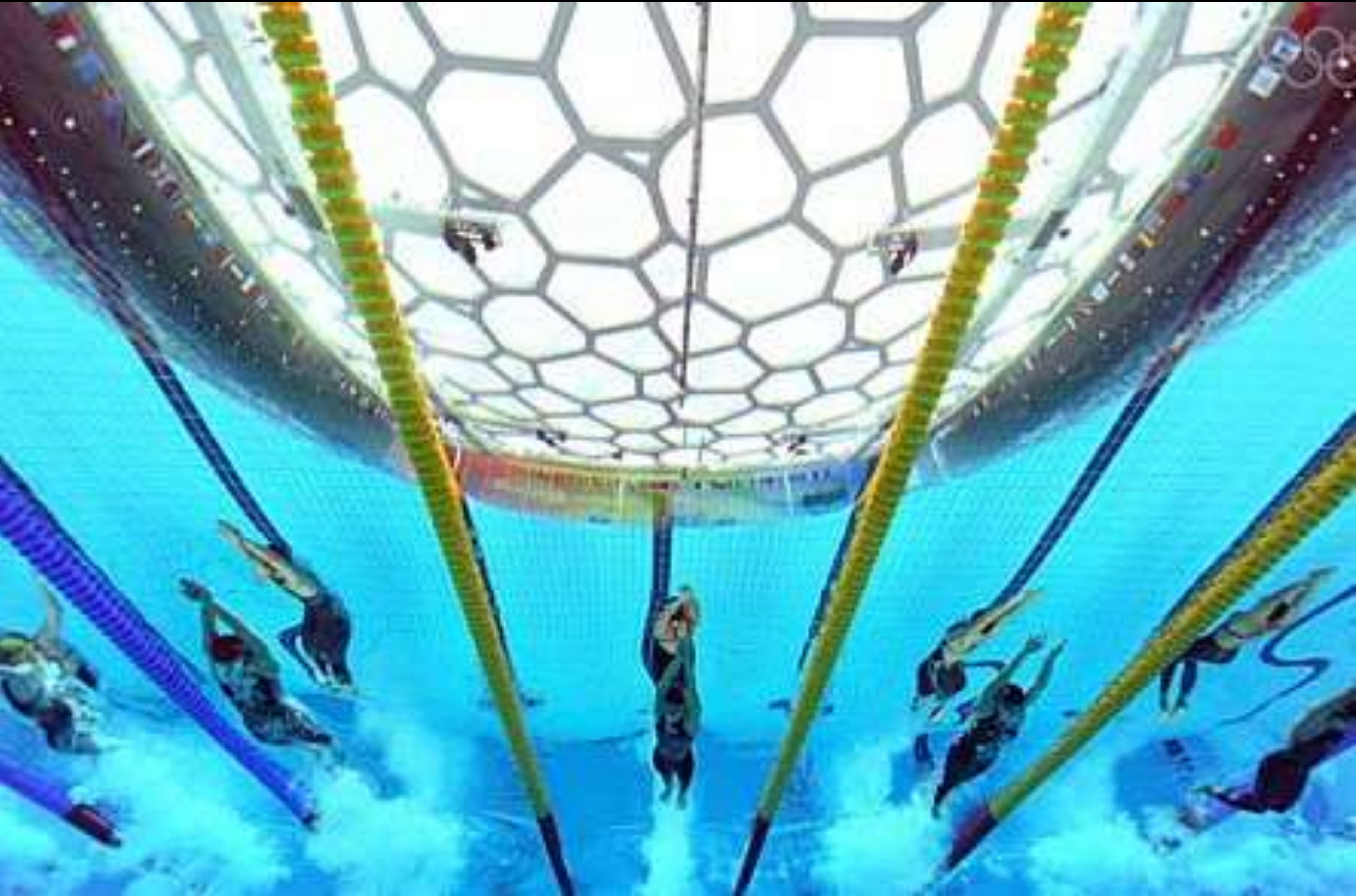
- **Akaike Information Criterion (AIC)** for model comparison
- **Mean absolute percentage error (MAPE)** to assess the relationship between estimated and observed performance
- **'Leave-one-out' cross-validation (LOO-CV)** for internal validation
- **Overfitting** of the model estimated from the ratio:

  RSS*(LOO-CV)–RSS(Apparent) / RSS(LOO-CV)

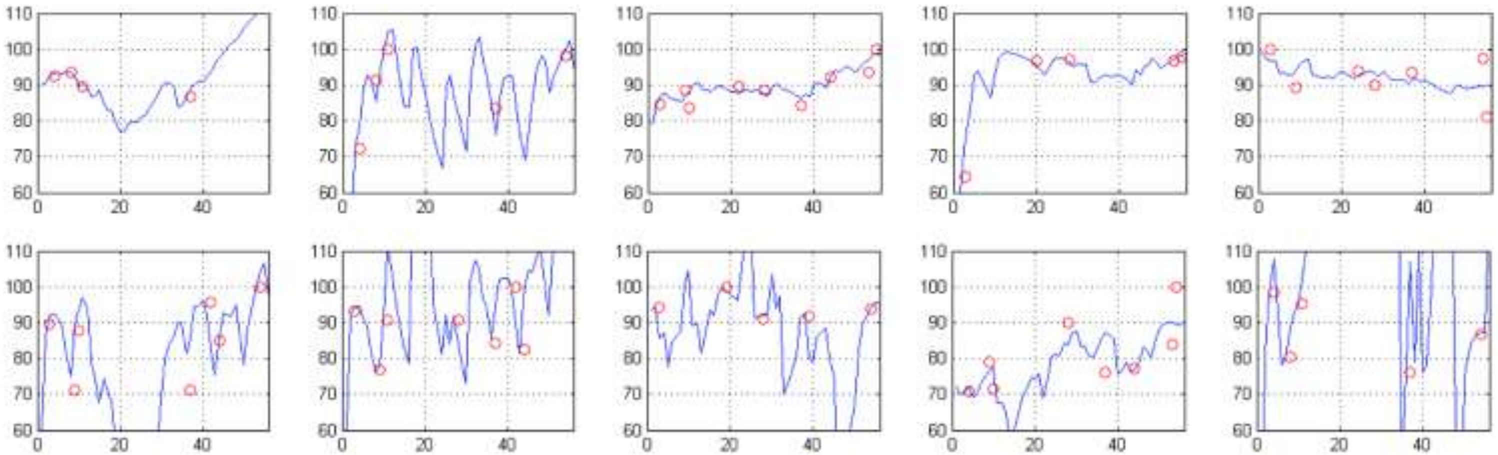
- Significance level: $P < 0.05$

*RSS = residual sum of squares

Results & discussion

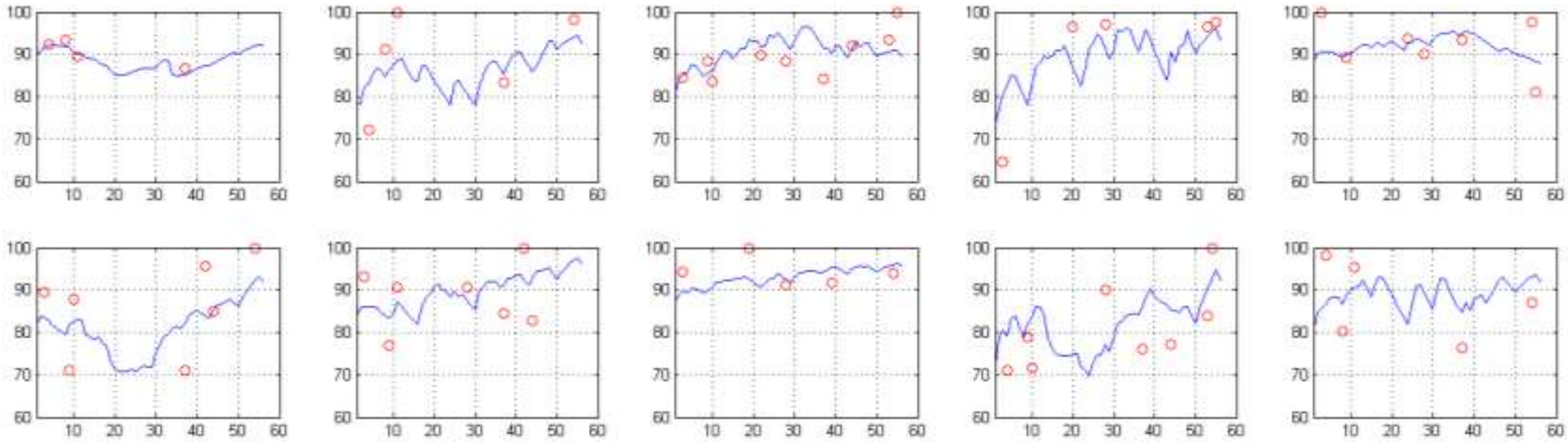


Model fitting (LinMod)



Subject	Tests	τ_1	τ_2	τ_3	p^*	k_1	k_2	r^2
1	4	60	18	6	90.1	0.013	-0.00002	1.000
2	5	11	15	2	40.4	0.285	-0.00066	1.000
3	9	27	30	3	79.0	0.067	-0.00018	0.779
4	5	11	7	10	50.0	0.134	-0.00015	1.000
5	7	19	30	1	99.0	0.041	-0.00016	0.193
6	7	6	7	2	56.6	0.25	-0.00043	0.678
7	7	8	9	2	57.7	0.436	-0.00097	0.982
8	5	5	5	1	93.6	0.309	-0.00151	0.921
9	8	11	3	2	72.3	0.024	-0.00017	0.556
10	5	54	24	4	26.4	0.603	-0.00157	1.000
Mean	**6.2**	**21.2**	**14.8**	**3.3**	**66.5**	**0.126**	**-0.00058**	**0.811**
SD	1.6	20.0	10.3	2.8	24.2	0.197	0.00058	0.269

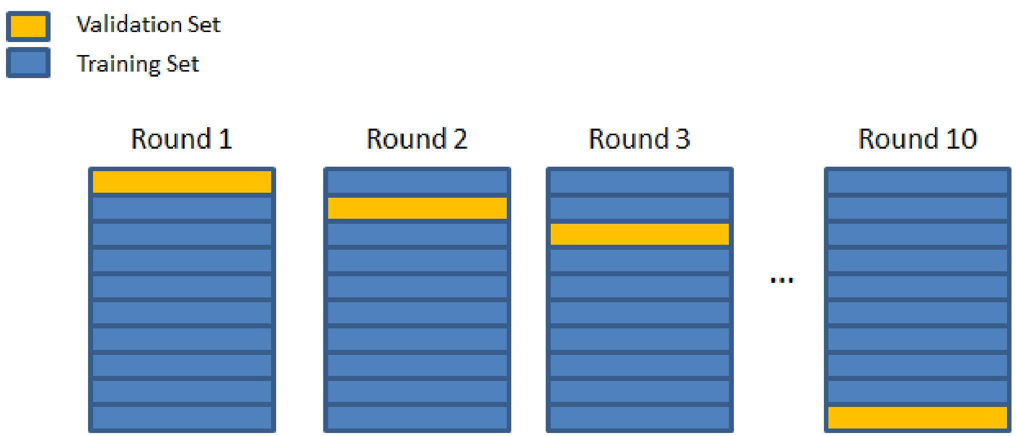
Model fitting (MixMod)



Subject	Tests	p^*	k_1	k_2	r^2
1	4	90.86653	0.00138	-0.43183	
2	5	78.81936	0.02137	-3.21892	
3	9	82.51912	0.01523	-2.36298	
4	5	67.37463	0.04036	-0.58666	
5	7	89.76855	0.00320	-0.68585	
6	7	79.08181	0.02093	-3.15820	
7	7	83.81238	0.01308	-2.06379	
8	5	87.48704	0.00698	-1.21366	
9	8	68.82400	0.03795	-5.53132	
10	5	88.23705	0.00574	-1.04015	
Mean	**6.2**	**90.86653**	**0.00138**	**-0.43183**	
SD	1.6	78.81936	0.02137	-3.21892	

Accuracy and internal validation

	MAPE (%)	RSS(Apparent)	RSS(LOO-CV)	Overfitting (%)
LinMod	**6.6**	1513	11982	**87.4**
MixMod	**6.8**	4998	6009	**16.8**



Conclusions

Conclusions

- The usual strategy for fitting the classical Busso-Banister model can be affected of an extreme overfitting
- This arises from the high number of observations per parameter needed (ca. 60 performance data points per individual)

Conclusions

- Another important limitation is that the adjustment is performed ignoring the correlation between the different observations on the same individual
- MixMod reduces overfitting because time parameters are common for all individuals (fixed-effects parameters) and are jointly estimated individual predictions (random-effects parameters)