



MODÉLISATION DES RÉPONSES AUX CHARGES D'ENTRAÎNEMENT ET GÉNÉRALISATION DES MODÈLES CHEZ DES SPORTIFS ÉLITES

Author: Frank Imbach¹ frank.imbach@umontpellier.fr
Co-auteur: Stephane Perrey, Romain Chailan, Thibaut Meline, Robin Candau¹ ; ²

¹ Seenovate, Montpellier, 34000, France

² EuroMov Digital Health in Motion, Univ Montpellier, IMT Mines Ales, Montpellier, 34090, France

Conférence Sciences²⁰²⁴: saison 2021

<https://sciences2024.polytechnique.fr>

May 28, 2021

Mots clefs: Modélisation, charge d'entraînement, généralisation, cross-validation.

1 Introduction

La relation entre la charge d'entraînement et ses effets sur la performance reste une énigme attendant d'être résolue, bien qu'étudiée depuis des décennies. Des premiers modèles décrivant les réponses aux charges d'entraînement sur la performance ont vu le jour il y a presque 50 ans avec notamment les modèles à réponse impulsionnelle, connus sous le nom de "Fitness-Fatigue" [4, 2]. Par la suite, des évolutions du modèle original ont été proposées avec notamment le modèle Dose-réponse variable [3], rendant l'état "Fatigue" fonction de l'accumulation des entraînements afin de mieux décrire les effets de l'entraînement sur la performance. Par ailleurs, des alternatives issues de la statistique ont été proposées, se différenciant principalement par une approche multivariée du problème [1]. Enfin, on note également l'emploi de modèles issus de l'apprentissage automatique semblant se démarquer lors de tâches de prédiction [5, 6] et étant pour la plupart construits au travers de procédures de validation.

A ce jour, aucun consensus relatif aux familles de modèles pour des tâches de prédiction de performances athlétiques ne semble néanmoins établi. De plus, la capacité de généralisation des modèles traditionnels reste à être explorée afin d'apporter de la crédibilité aux interprétations physiologiques et pratiques issues des paramètres des modèles. Au travers de cette étude, nous avons donc proposé une méthodologie robuste et transférable basée sur la généralisation des modèles dans un contexte de modélisation de la performance athlétique.

2 Méthode utilisée

1. Dispositif expérimental

Pendant 3 mois, nous avons collecté des données d'entraînement et de performance auprès de sept membres de l'équipe de France de Short-track.

Les participants réalisaient des contre-la-montre avec départ arrêté plusieurs fois par semaine. Au total, 248 performances ont été enregistrées pour l'ensemble du groupe d'athlètes. Les charges d'entraînement brutes ont été calculées à partir de types d'entraînement variés et déterminées par des estimations physiques objectives (en accord avec les variations d'énergie cinétique), physiologiques (pondération selon les cinétiques de fréquence cardiaque et de concentration en lactate sanguin à l'effort) et subjectives (perception de l'effort par l'athlète). Par ailleurs, des paramètres inhérents à la programmation et à l'environnement tels que les durées de récupération inter-sessions et la qualité de la glace furent considérés.

2. Méthode numérique

Afin d'associer les informations relatives à l'historique des entraînements à une performance particulière, les charges d'entraînement ainsi que les paramètres d'efforts journaliers ont été agrégés sur la base d'une réponse impulsionnelle et d'une réponse sérielle bi-exponentielle, à l'image d'adaptations physiologiques de l'organisme en réponse à l'effort [7].

De plus, l'information temporelle au travers de performances autocorrélées présentant un pouvoir explicatif, la performance précédente était ainsi considérée comme prédicteur de la performance du jour.

A des fins de comparaison, chaque modèle a bénéficié d'une validation par validation croisée (procédure détaillée sur la Figure 1). Du fait de variables non indépendantes et identiquement distribuées, la validation des modèles a été effectuée sur des échantillons respectant l'ordre temporel des données.

3. Méthode théorique

A partir de données d'entraînement agrégées selon des principes physiologiques avérés et de données

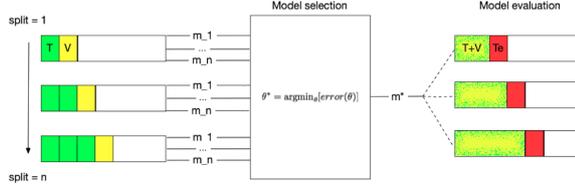


Figure 1: Schéma de la validation croisée de série-temporelle suivi d'une d'évaluation finale sur des échantillons de données non biaisés. Les lettres T, V et Te représentent les échantillons de données utilisés pour l'entraînement, la validation et l'évaluation respectivement.

brutes relatives aux facteurs techniques et environnementaux, nous avons comparé la capacité de généralisation des modèles construits dans une procédure de validation croisée. Par ailleurs, du fait d'une faible taille d'échantillon, chaque modèle comparé au modèle Dose-réponse (DR_I) considéré ici comme référence [3] était construit soit par athlète (M_I), soit sur l'ensemble du groupe d'athlètes (M_G). Une sélection de modèle opérée durant la procédure de validation croisée devrait permettre à la fois de garantir une généralisation optimale et d'en évaluer ses capacités lors de futures tâches de prédictions, quel qu'en soit le contexte et la discipline sportive étudiée.

3 Résultats obtenus

Les plus grandes capacités de généralisation des modèles ont été attribuées aux régressions régularisées Elastic net ($ENET_I$, $ENET_G$) et aux régressions en composantes principales (PCR_I , PCR_G), en comparaison au modèle de référence DR_I . Par ailleurs, à l'exception de PCR_G , chaque modèle a démontré une qualité prédictive supérieure au modèle DR_I . Une classification de ces résultats est présentée dans en Table 1.

Table 1: Détail des comparaisons par paires pour les capacités de généralisation et les qualités prédictives. β_{diff} représente la différence dans les estimations des paramètres entre le modèle DR_I et sa comparaison.

Comparison	β_{diff}	CI_{lower}	CI_{upper}	$Pr(> t)$	Criterion
$DR_I - ENET_G$	0.057	0.045	0.069	< 0.001	Generalisation
$DR_I - PCR_G$	0.032	0.020	0.044	< 0.001	Generalisation
$DR_I - PCR_I$	0.031	0.010	0.053	0.005	Generalisation
$DR_I - ENET_I$	0.027	0.007	0.050	0.012	Generalisation
$DR_I - RF_I$	-0.023	-0.045	-0.001	0.032	Generalisation
$DR_I - RF_G$	-0.026	-0.038	-0.014	< 0.001	Generalisation
$DR_I - ENET_G$	0.032	0.016	0.048	< 0.001	Prediction
$DR_I - ENET_I$	0.025	0.004	0.047	0.020	Prediction
$DR_I - RF_I$	0.022	0.001	0.044	0.043	Prediction
$DR_I - PCR_I$	0.020	-0.002	0.041	0.072	Prediction
$DR_I - RF_G$	0.014	-0.002	0.029	0.088	Prediction
$DR_I - PCR_G$	-0.051	-0.067	-0.035	< 0.001	Prediction

Une application des modèles proposant les meilleures capacités de généralisation au regard du modèle de référence a permis d'illustrer la sensibilité des modèles pour un athlète représentatif (Figure 2). Les modèles $ENET_G$ et PCR_G semblent davantage capter les vari-

ations de performances que le modèle de référence.

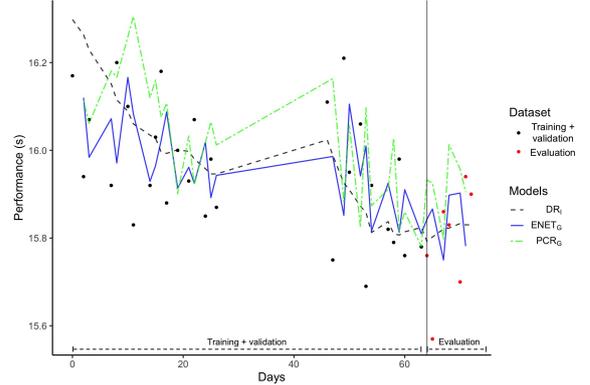


Figure 2: Performance des modèles pour un athlète représentatif. Les lignes continues et pointillées représentent le modèle DR_I et les deux modèles offrant la plus grande généralisation. Sur cet exemple, les données d'entraînement (80% des données totales, combinant les échantillons d'entraînement et de validation) et d'évaluation (20% des données totales, l'échantillon d'évaluation) sont séparées par la ligne verticale continue.

4 Modèles

Le modèle de référence noté DR_I a été comparé à des fins de généralisation et de précision lors de prédictions. Il se définit comme un modèle linéaire avec paramètre variable au cours du temps, où $w(l)$ est une fonction discrète représentant une charge d'entraînement journalière (Equation 1).

En comparaison, deux procédures de régularisation sont utilisées à savoir une régression Elastic net et une régression en composante principale, ainsi qu'une régression par forêts aléatoires prenant toutes en entrée les variables indépendantes sous forme de données brutes ou de données agrégées. Les réponses impulsives et sérielles bi-exponentielles associées aux fonctions de transfert pour la construction d'agrégations sont présentés par les Equations 2, 3.

4.1 Equations

$$\hat{y}^{(\text{busso})}(t) = y^* + k_1 \sum_{l=1}^{t-1} w(l) e^{-\frac{(t-l)}{\tau_1}} - \sum_{l=1}^{t-1} k_2(l) w(l) e^{-\frac{(t-l)}{\tau_2}}, \quad (1)$$

$$\text{avec } k_2(l) = k_3 \sum_{m=1}^l w(m) e^{-\frac{(l-m)}{\tau_3}}.$$

$$X_n(t) = \sum_{l=1}^t w(l) \left(e^{-\frac{(t-l)}{\tau_r}} \right) \quad (2)$$

$$X_n(t) = \sum_{l=1}^t w(l) \left(\left(1 - e^{-\frac{(t-l)}{\tau_G}} \right) U + e^{-\frac{(t-TD-l)}{\tau_D}} | U - 1 | \right) \quad (3)$$

$$\text{avec } U = \begin{cases} 1 & \text{si } t < TD \\ 0 & \text{autrement.} \end{cases}$$

5 Conclusion et perspectives

Dans cette étude, nous avons proposé une méthodologie transférable basée sur l'évaluation de la capacité de généralisation des modèles dans un contexte de modélisation de performances athlétiques. Chacun des modèles utilisés ont été soumis à une validation croisée dans le respect de la structure temporelle des données. Ainsi, les regressions régularisées ont offert une plus grande généralisation que le modèle dose-réponse, pris pour référence. Par ailleurs, l'écart de généralisation et les qualités prédictives des modèles se justifient en partie par une approche multivariée du problème, où la performance athlétique dépend d'un certain nombre de facteurs propres à l'entraînement et à son environnement.

Sur cette base, d'autres applications seraient envisageables avec notamment des simulations conditionnelles des variables indépendantes en bénéficiant des relations connues entre certains paramètres d'effort, qu'ils soient de nature objective ou subjective.

6 Remerciements

Nous tenons à remercier la Fédération Française des Sports de Glace, l'Institut National du Sport, de l'Expertise et de la Performance, Dr. Anthony MJ Sanchez et Robert Solsona (Laboratoire Européen Performance Santé Altitude, University of Perpignan Via Domitia) pour leur collaboration.

References

- [1] Marta Avalos, Philippe Hellard, and Jean-Claude Chatard. Modeling the training-performance relationship using a mixed model in elite swimmers. *Med. Sci. Sports Exerc.*, 35(5):838, 2003.
- [2] EW Banister and TW Calvert. Planning for future performance: implications for long term training. *Can. J. Appl. Sport Sci.*, 5(3):170–176, 1980.

- [3] Thierry Busso. Variable dose-response relationship between exercise training and performance. *Med. Sci. Sports Exerc.*, 35(7):1188–1195, 2003.
- [4] Thomas W Calvert, Eric W Banister, Margaret V Savage, and Tim Bach. A systems model of the effects of training on physical performance. *IEEE Trans. Syst. Man. Cybern. SMC*, 6(2):94–102, 1976.
- [5] Jürgen Edelmann-Nusser, Andreas Hohmann, and Bernd Henneberg. Modeling and prediction of competitive performance in swimming upon neural networks. *Eur. J. Sport. Sci.*, 2(2):1–10, 2002.
- [6] Lachlan JG Mitchell, Ben Rattray, John Fowlie, Philo U Saunders, and David B Pyne. The impact of different training load quantification and modelling methodologies on performance predictions in elite swimmers. *Eur. J. Sport Sci.*, pages 1–10, 2020.
- [7] Antony G Philippe, Fabio Borrani, Anthony MJ Sanchez, Guillaume Py, and Robin Candau. Modelling performance and skeletal muscle adaptations with exponential growth functions during resistance training. *J. Sports Sci.*, 37(3):254–261, 2019.