Modelling responses to training loads to lead athletic performance in sports

Frank Imbach

▶ **To cite this version:**

Frank Imbach. Modelling responses to training loads to lead athletic performance in sports. Computer Science [cs]. Université de montpellier, 2021. English. NNT : . tel-03547563

HAL Id: tel-03547563
https://hal.science/tel-03547563

Submitted on 28 Jan 2022

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

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

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Sciences du Mouvement Humain

École doctorale 463 - SMH

Unité de recherche EuroMov Digital Health in Motion

Modelling responses to training loads to lead athletic performance in sports

Présentée par Frank IMBACH
Le 09 décembre 2021

Sous la direction de Stéphane PERREY
et Robin CANDAU

Devant le jury composé de

Pierre DRUILHET, Pr, Laboratoire de Mathématiques Blaise Pascal, UMR CNRS 6620, Université Clermont Auvergne, Aubière, France	Rapporteur
Thierry BUSSO, Pr, Laboratory of Exercise Physiology, Jean Monnet University of Saint-Etienne, PRES Lyon, Saint-Etienne, France	Rapporteur
Christophe LEY, Pr, University of Luxembourg - Department of Mathematics, Luxembourg	Examinateur
Anne LAURENT, Pr, LIRMM, Univ Montpellier, Montpellier, France	Examinatrice
Jacques PRIOUX, Pr, Faculty of sport science, Movement, Sport and Health laboratory, ENS Rennes, Rennes, France	Président, examinateur
Stéphane PERREY, Pr, EuroMov Digital Health in Motion, Univ Montpellier, IMT Mines Ales, Montpellier, France	Directeur de thèse
Robin CANDAU, Pr, DMeM, Univ Montpellier, INRAe, Montpellier, France	Co-Directeur de thèse
Romain CHAILAN, Seenovate, Montpellier, France	Invité

UNIVERSITÉ DE MONTPELLIER

Affidavit

I, undersigned, Frank Imbach, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific direction of Stéphane Perrey and Robin Candau, in accordance with the principles of honesty, integrity and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with both the french national charter for Research Integrity and the University of Montpellier charter on the fight against plagiarism.

This work has not been submitted previously either in this country or in another country in the same or in a similar version to any other examination body.

Saturday, 09 October 2021

Cette œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

List of publications and conference participation

List of publications produced through the thesis:

1. **Imbach, F.**, Sutton-Charani, N., Montmain, J., Candau, R., & Perrey, S. (2021). The use of Fitness-Fatigue models for sport performance modelling : conceptual issues and contributions from machine-learning (in revision)
2. **Imbach, F.**, Perrey, S., Chailan, R., Méline, T., & Candau, R. (2022). Training load responses modelling and model generalisation in elite sports. *Scientific reports*. doi : https ://doi.org/10.1038/s41598-022-05392-8.
3. **Imbach, F.**, Candau, R., Chailan, R., & Perrey, S. (2020). Validity of the Stryd power meter in measuring running parameters at submaximal speeds. *Sports*, 8(7), 103.

Participation in conferences during the thesis:

1. **Imbach, F.**, Perrey, S., and Candau, R. (28 October 2021). A new model of training load quantification in resistance training. **19th ACAPS international congress**, Oral communication, Montpellier, France.
2. **Imbach, F.**, Perrey, S., Chailan, R., Méline, T., and Candau, R. (28 May 2021). Training load responses modelling and model generalisation in elite sports. **Sciences2024**, Oral communication, France.
3. **Imbach, F.**, Perrey, S., Chailan, R., Méline, T., and Candau, R. (28 October 2020). Training load responses modelling in elite sports : how to deal with generalisation ? **European Congress of Sport Science international congress**, Oral communication, Seville, Spain.
4. **Imbach, F.**, Perrey, S., and Candau, R. (04 November 2019). La data au service de la performance sportive, **ANRT**, Oral communication, Paris, France. In : Les sciences au service de la performance sportive, Association nationale de la recherche et de la technologie, p. 189–199.
5. **Imbach, F.**, Perrey, S., Chailan, R., Méline, T., and Candau, R. (20 October 2019). Modelling responses to training loads in short track speed-skating. **18th ACAPS international congress**, Oral communication, Paris, France.

Résumé

Les premiers modèles des effets de l'entrainement sur la performance athlétique sont connus sous le nom de modèles "Fitness-Fatigue" (FFM). Un inconvénient majeur des FFMs réside dans le fait qu'ils ne sont constitués que d'une seule donnée d'entrée, bien que la performance athlétique soit multifactorielle. Dès lors, des approches multivariées propres aux statistiques et à l'apprentissage automatique ont été proposées pour différentes applications sportives.

La quantification de la charge d'entraînement (CE) pour l'entraînement en résistance constitue une problématique de recherche à part entière. Dans une première étude, nous avons proposé une nouvelle méthode de quantification de la CE en accord avec des réponses physiologiques multiples à l'effort. Après avoir modélisé les profils couple-vitesse chez les participants, nous avons évalué des réponses physiologiques lors de 3 sessions d'effort en résistance à intensités variées et de volume égal. Les hautes intensités ont entrainé une fatigue musculaire plus importante caractérisée par des altérations neuromusculaires. A l'inverse, la consommation d'oxygène ainsi que les modifications métaboliques étaient supérieures lors d'efforts de plus faible intensité, indiquant des contributions énergétiques différentes. Ainsi, nous avons proposé un indice de CE basé sur les altérations neuromusculaires observées à l'effort. Pondérer exponentiellement la CE par une constante de décroissance du taux d'augmentation de la force a montré de plus grandes corrélations avec les réponses physiologiques étudiées. Par ailleurs, l'information compressée de données multivariées dans une seule composante suite à analyse en composantes principales pourrait représenter un indice de CE.

Dans la deuxième étude, nous avons proposé une méthodologie de modélisation basée sur la généralisation des modèles. Nous avons comparé un modèle dose-réponse à des procédures de régularisation et modèles d'apprentissage automatique multivariés chez des patineurs élites. Les modèles de régularisation ont montré de meilleures performances en termes de généralisation et de précision. De plus, des modèles construits sur l'ensemble du groupe plutôt que par athlète apparaissaient plus pertinents dans un contexte d'échantillons de petite taille. Enfin, des approches en apprentissage automatique telles que les méthodes d'apprentissage ensemblistes pourraient améliorer le pouvoir prédictif des FFMs.

Dans la troisième étude, nous avons modélisé les profils accélération-vitesse à partir des mesures de systèmes de localisation par satellite (GPS), puis avons tenté

de prédire les coefficients de la relation accélération-vitesse. Tout d'abord, des modèles de prévision de séries chronologiques ont constitué une base de modélisation. Nous les avons par la suite comparés à une régression linéaire régularisée et un réseau de neurones récurrents utilisant des variables propres au GPS. Enfin, nous avons extrait des variables directement à partir des données GPS brutes pour effectuer la tâche de prédiction. Aucune différence significative n'a été observée entre les modèles en termes de précision. Étant donné le caractère multifactoriel de la performance athlétique, les performances prédictives étaient acceptables. L'utilisation de données extraites des domaines temporels et fréquentiels à partir des signaux bruts a montré des performances comparables aux autres modèles. Les données brutes semblent donc avoir un intérêt et devraient être analysées pour des problématiques relatives à la performance athlétique et à la survenue de blessure.

Enfin, nous avons développé un système de suivi de l'entrainement pour des coureurs de fond. L'application propose un module de suivi de l'entrainement et un modèle prédictif basé sur une modélisation physiologique de la performance en course à pied. Un second développement a été réalisé sous la solution SAP analytics cloud, produisant des rapports automatisés et un suivi de joueurs d'une équipe de Rugby.

Abstract

The first models of training effects on athletic performance emerged with the work of Banister and Calvert through the so-called Fitness-Fatigue model (FFM). One major drawback of FFMs is that the features stem from a single source of data. That is not in line with the existing consensus about a multifactorial aspect of athletic performance. Hence, multivariate modelling approaches from statistics and machine-learning (ML) emerged.

A research issue arises from the quantification of training Loads (TL) in resistance training (RT) which lack of physiological evidence. In the first study, we provided a new method of TL quantification in RT based on physiological observations. To achieve that, we initially modelled the torque-velocity profiles of fifteen participants during an isokinetic leg extension task and assessed a set of physiological responses to various resistance exercises intensities. Each session was volume-equated according to the formulation of volume load (i.e. the product of the number of repetitions and the relative intensity).

Higher led to greater muscular fatigue described by neuromuscular impairments. Conversely, systemic and local pulmonary responses (measured through oxygen uptake) and metabolic changes (according to blood lactate concentrations) were more significant at low intensities, suggesting different contributions of metabolic pathways.

From these results, we provided a new index of TL based on the neuromuscular impairments observed at exercise. We showed that to exponentially weight TL by the average rate decay of force development rate yielded better correlations with any of the significant physiological responses to exercise. In addition, information compressed within a principal component could be a valuable TL index.

In the second study, we provided a robust modelling methodology that relies on model generalisation. Using data from elite speed skaters, we compared a dose-response model to regularisation methods and machine-learning models.

Regularisation procedures provided the greatest performances in both generalisation and accuracy. Also, we highlighted the pertinence of computing one model over the group of athletes instead of a model per athlete in a context of a small sample size.

Finally, ML approaches could be a way of improving FFMs through ensemble learning methods.

In the third study, we modelled acceleration-velocity directly from global positioning system (GPS) measurements and attempted to predict the coefficients of the relationship between acceleration and velocity.

First, a baseline model was defined by time-series forecasting using game data only. Then, we proceeded to multivariate modelling using commercial features. A regularised linear regression and a long short term memory neural network were compared. Finally, we extracted features directly from raw GPS data and compared these features to the commercial ones for prediction purposes.

The results showed only slight differences between model accuracy, and no models significantly outperformed the baseline in the prediction task. Given the multifactorial nature of athletic performance, using only GPS data for predicting such athletic performance criterion provided an acceptable accuracy. Using time-domain and frequency-domain features extracted from raw data led to similar performances compared to the commercial ones, despite being evidence-based. It suggests that raw data should be considered for future athletic performance and injury occurrence analysis.

Lastly, we developed an athlete management system for long-distance runners. This application provided an athlete monitoring module and a predictive module based on a physiological model of running performance.

A second development was realised under the SAP analytics cloud solution. Team management and automated dashboards were provided herein, in close collaboration with a professional Rugby team.

Acknowledgements

To my colleagues at Seenovate
Thank you for the time spent together, the push up challenges, the funny lunches in the MIBI's open space and, of course, your assistance during these three years!

To *Jean-Michel* – thank you for your trust and for allowing me to complete the thesis.

To *Romain* – I give you a special thank. You have led me all along the thesis and made computer science another (nth I would say) domain of particular interest. Last but not least, I am delighted to have attended your first-ever and last 313 landings! You will remain the lord of 312 for decades for sure, be proud!

To my colleagues at EuroMov
Thank you for the time spent together, the raclette band will be forever etched in my memory.

Pauline, Clementine and Louis – thank you friends, faluches or écolos, I can't wait for the next whine/cheese/sausage apéro !

To my supervisors
Stéphane, I am still impressed by your wit, rigour, and research skills, evidenced by your prodigious academic contribution to exercise physiology and neurophysiology fields. Since the courses at University, I was convinced that you would be the director I would need for an outstanding supervision. I was not wrong and have learnt a lot from your side.

Robin, PhD students should all have a supervisor like you. Your cleverness, caring, and exceptional involvement with your students make you a real professor/friend to rely on. I remember our first call, you taught me how to get the fellowship with a very first rehearsal while skiing/climbing the Alps. You made it all possible since that day, thank you!

Both of you are world-class researchers, and I could not hope for better supervision. You have been an inspiration throughout these three years, and I look forward to working with you in the future.

To my family
Dad, you have given me a love for sports when I was a child. You always supported me in my scholarship and professional projects. I could not thank you enough for that and will always be happy to share a bike ride or windsurfing session with you!

Mamá, siempre has creído en mí, ¿quién iba a pensar que veníamos hasta aquí? Creo que hemos recorrido un largo camino desde el instituto, ¡pero por fin me siento realizada! La vida continúa y nos tiene reservadas grandes y bellas cosas.

I hope I will continue making you proud!

Margot, merci de m'avoir supporté dans tous mes projets, et recadré quand cela s'avérait nécessaire ! Tu as été ma force et sans toi je ne serais évidemment pas arrivé jusqu'ici. Tu peux être fière de toi, et de notre magnifique petite famille !

Antone, Brune, vous êtes arrivés en m'apportant un immense bonheur pendant ces trois années, rendant ce travail parfois laborieux si plaisant !

Michel, beau papa, je te dédie cette thèse ainsi que tout ce que j'ai pu accomplir jusqu'ici. Depuis mes 5 ans, tu as été un pillier dans ma vie et a fait ce que je suis aujourd'hui. Je continuerai chaque jour de m'inspirer de ta personne, certainement la plus belle qu'il fut dans ce monde.

Contents

Affidavit 2

List of publications and conference participation 3

Résumé 4

Abstract 6

Acknowledgements 8

Contents 10

List of Figures 13

List of Tables 19

List of acronyms 21

Glossary 30

Introduction 31

1 State of the art 36
- 1.1 Training load, definition and quantification . . . 36
  - 1.1.1 Objective measures . . . 37
  - 1.1.2 Subjective measures . . . 54
- 1.2 Modelling the effects of training . . . 59
  - 1.2.1 A collection of Fitness-Fatigue models . . . 59
  - 1.2.2 Analogies with the fitness-fatigue framework . . . 75
  - 1.2.3 Contributions from statistics and computer science . . . 81
- 1.3 Conclusion . . . 86

2 Training loads: from an objective training-load quantification to physiological responses 87
- 2.1 Resistance training: from an objective training-load quantification to physiological responses . . . 88
  - 2.1.1 Introduction . . . 88

2.1.2 Experimental setup . . . 89
2.1.3 Data collection . . . 94
2.1.4 Physiological responses to resistance exercises . . . 101
2.2 Towards a new model of training effect quantification . . . 125
2.2.1 A physiological model of training load quantification in resistance training . . . 126
2.2.2 A linear combination of quantification methods and exercise related variables . . . 128
2.2.3 Relationship between training load quantification methods and physiological responses . . . 133
2.2.4 Conclusion . . . 134
2.2.5 Take-home message . . . 136

3 Modelling responses to training loads and athletic performance prediction 137
3.1 Generalities . . . 138
3.1.1 Context . . . 138
3.1.2 Model selection and evaluation . . . 139
3.2 Training load responses modelling and model generalisation: application in short-track speed skating . . . 144
3.2.1 Data . . . 144
3.2.2 Modelling Methodology . . . 149
3.2.3 Evaluation of model performances . . . 152
3.3 Fitness-fatigue models: advantages, conceptual issues and contribution from machine-learning . . . 160
3.3.1 Making the most of control theory . . . 160
3.3.2 Current Fitness-fatigue framework and related issues . . . 161
3.3.3 A machine-learning perspective of the problem . . . 163
3.3.4 Take home message . . . 164
3.4 Using global positioning systems for modelling athletic performances 166
3.4.1 Methodological approach . . . 167
3.4.2 Predicting the acceleration-velocity profile from GPS summarised features . . . 173
3.4.3 Conclusion . . . 187
3.4.4 Take home message . . . 188

4 A decision support system for team and athletes 190
4.1 Introduction . . . 190
4.2 Validation of a running power meter . . . 192
4.2.1 Methodological approach . . . 192
4.2.2 Verification of reference measures . . . 197
4.2.3 Stryd and reference measures comparisons . . . 200
4.2.4 Conclusion . . . 207

4.2.5 Take-home message 208
4.3 *RunningPerfApp*: A *shiny* web-application for long distance runners 208
4.3.1 Monitoring athlete status 208
4.4 Team management and data visualisation with SAP analytics cloud 211
4.4.1 Monitoring team and players body mass 212
4.4.2 Monitoring player performances 214
4.4.3 Monitoring training from global positioning system data . . 215
4.4.4 Monitoring Well-being 215

Conclusion **217**

Bibliography **218**

APPENDIX **251**
A Appendix: Derivation of Equation 1.13 251
B Derivation of Equations 1.33 to 1.35 252
C Mathematical formulation of fitness-fatigue model combined to a kalman filter, according to Kolossa, Azhar, Rasche, et al. 2017 . . . 254
D Discretisation of model equations, taken from Busso 2017 256
E Hierarchical mixed models used in Section 2.1 257
F Appendix: Training load quantification in Short track speed-skating 259
G Appendix: Fitted parameters using elastic net regularisation and principal component regression 261
G.1 Elastic net regularisation 261
G.2 Principal component regression: contribution of variables per dimension 262
H Appendix: Fitted parameters using elastic net regularisation and principal component regression 263
H.1 Reference Measures, section 4.2.2 263
H.2 Comparisons between Stryd and reference measures 264

List of Figures

1.1 Excess post-exercise oxygen consumption (EPOC) after exhaustive submaximal exercise (71-80 minutes at 69-78% of $\dot{V}O_2max$), taken from Børsheim and Bahr 2003. Circles represent the recorded $\dot{V}O_2$ values and the solid line shows the prolonged component. 44

1.2 Schematic representation about the use of fitness-fatigue model (FFM) for making predictions and determining the optimal command , taken from Busso and Thomas 2006. 61

1.3 Diagram representation of the two component Fitness-Fatigue model, taken from Morton, Fitz-Clarke, and Banister 1990 62

1.4 Block diagram of a Kalman filter (KF) system, taken from Kolossa, Azhar, Rasche, et al. 2017. $\mathbf{B_{k-1}}$ is the time-varying input matrix that also contains the two exponential decays on day $k-1$; $\mathbf{C}$ is the output matrix that shows the influence of each state variables on the measured performance; $\mathbf{K_k}$ denotes the optimal kalman gain. . . 74

1.5 Diagram of the Performance Potential (PerPot) model, adapted from Perl 2001. The terms DSO, DS and DR denote the delays of strain overflow, delays of strain flow and delays of response flow, respectively. 76

1.6 Block diagrams of four secondary signal models, reprinted from Busso 2017. 80

2.1 Figures 2.1a, 2.1b and 2.1c represent torque-velocity (T-V) profiles modelled from an isokinetic knee extensions at different velocities. Values are normalised to the maximal torque measured at 0.524 rad s^{-1}. Red dots represent the measured torque for each velocity, the line represents the modelled T-V profile according to Equation 2.1. Figure 2.1d shows the distribution of decay rate constants (unit of normalised torque per rad s^{-1}, according to Equation 2.1). 91

2.2 Model of relationship between relative intensity in percentage of repetition maximum (RM) and the theoretical number of repetitions that can be achieved, according to Reynolds, Gordon, and Robergs 2006 and Equation 2.2. Red asterisks represent the retained values for the relative intensity and the number of repetition included in the protocol. 93

2.3 oxygen uptake ($ml.min^{-1}$) ($\dot{V}O_2$) recorded from a participant during resistance exercise protocol: moderate intensity session ($C2$). 95

2.4 Changes in oxygenation index, the difference between oxyhemoglobin and deoxyhemoglobin concentration ($\Delta\,[\mathrm{Hb_{diff}}]$) over time during the three testing conditions resistance exercise protocol: low intensity session ($C1$), $C2$, and resistance exercise protocol: heavy intensity session ($C3$) for (a), (b), and (c), respectively. 98
2.5 Power spectral density of vastus lateralis (VLat) calculated from short-time Fourier transformations (STFT) over four samples (from the second, fifteenth, thirteenth and forty-fifth samples of a single repetition). 100
2.6 Picture of a participant performing a testing session with all associated measurements. 101
2.7 Distributions of changes in serum plasma cortisol concentrations ($[cort_p]$) after having completed the three testing sessions. The asterisk denotes the level of significance $p < 0.05$ for differences between distributions. 103
2.8 Distribution of Heart rate (HR) slopes and HR amplitudes across testing conditions. The asterisks denote the level of significance $p < 0.05$ for differences between distributions. 105
2.9 Distribution of (a) session-averaged $\dot{\mathrm{V}}\mathrm{O}_2$ slopes computed over recovery phases and (b) session-averaged $\dot{\mathrm{V}}\mathrm{O}_2$ amplitudes. The asterisks * denote the level of significance $p < 0.05$ for differences between distributions. 107
2.10 Distribution of (a) energy expenditure from exercise phases only, and (b) energy expenditure from exercise and post-exercise revcovery phases. The asterisks *** denote the level of significance $p < 0.001$ for differences between distributions. 107
2.11 Distribution of (a) $\Delta\,[\mathrm{Hb_{diff}}]$ slopes and (b) tissue saturation index (percentage of ratios of absorbance) (TSI) slopes at exercise. Asterisks denote the level of significance for $*$, $**$, and $***$ being equivalent to $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. 109
2.12 Changes in $\Delta\,[\mathrm{Hb_{diff}}]$ over three isokinetic leg extensions performed at $0,174$ rad s^{-1}. For each repetition, short events of ischemia - perfusion phenomena are represented by squares. The slope of $\Delta\,[\mathrm{Hb_{diff}}]$ denoted α is computed over the longest –unbiased– interval of the first repetition. 110
2.13 Distribution of (a) $\Delta\,[\mathrm{Hb_{diff}}]$ slopes and (b) TSI slopes during recovery phases. 111
2.14 Distribution of change in total hemoglobin concentration, the sum of oxyhemoglobin and deoxyhemoglobin concentration ($\Delta\,[\mathrm{tHb}]$) slopes during recovery phases. 112

2.15 Distribution of total mechanical work (J) (W_{mech}), normalised averaged torque, maximal rate of force development computed from 20ms time intervals($Nm.s^{-1}$) (RFD_{peak}), rate of force development from the onset to 100ms of exercise ($Nm.s^{-1}$) (RFD_{0-100}) and the area under the torque curve ($Nm.s$) (impulsion) measures across testing sessions. Asterisks denote the significance level (*** being related to $p < 0.001$). In (c), (d) and (e), the distribution of the third session is significantly different from the first session (brackets are not displayed for clarity). 114
2.16 Distribution of regression slopes for changes in (a) RFD_{peak} and (b) RFD_{0-100} across knee extension repetitions. 116
2.17 Electromyographic signals of the VLat over two sets, recorded during $C2$ for a representative participant. The thin line represents the averaged signal surrounded by a one standard deviation ribbon and displayed over all repetitions. 119
2.18 Distributions of summated Surface electromyography (EMG) signals from leg extensors and normalised averaged torque over repetitions. 122
2.19 Distribution of regression slopes for changes in median frequency (MDF) across repetitions of isokinetic knee extensions, with (a) changes over conditions and (b) by muscle changes. 124
2.20 Representation of (a) the non-linear relationship between the rate decay of RFD_{peak} and the relative intensity, and (b) the relationship between a neuromuscular breakdown given by the ratio $\frac{1}{f(x)}$ and the relative intensity. 128
2.21 Circle of correlations with (a) the quality of each variable representation (cos^2) and (b) the circle of individuals. In (a), variables in blue are illustrative variables, not accounted for in the calculation of distance between individuals. In (b), groups 1, 2 and 3 represents the testing sessions $C1$, $C2$ and $C3$, respectively. 131
2.22 (a) Representation of principal components and (b) $\cos^2$ of variables to the first dimension. 132

3.1 Diagram of leave-one-out cross-validation. Green and red circles represent observations used for model training and validation, respectively. 140
3.2 Diagram of leave-p-out cross-validation, with $p = 3$. Green and red circles represent observations used for model training and validation, respectively. 140
3.3 Diagram of K-fold cross-validation, with $k = 5$. Green and red squares represent subsets of data (i.e. folds) used for model training and validation, respectively. 141

3.4 Model selection according to prequential approaches. (a) represents a time-series block cross-validation (CV) with a growing window, (b) represents a variant with a sliding window and (c) a variant with of a gap introduced within the sliding window approach. 143
3.5 Cumulative daily training loads of a representative athlete following (a) the impulse response function (X_2, Equation 3.1) and (b) the serial bi-exponential response function (X_3, Equation 3.2). (c) illustrates the raw daily training loads X_1, expressed by w(t). In (a) and (b), dots represent daily values of the cumulative training load and vertical solid lines indicate occurrence of training sessions. Values are represented in arbitrary units (a.u). 147
3.6 Distributions of models' performance. (a) shows root mean square error (RMSE) distributions of each individual models and (b) the models computed on the whole group. Within boxplot midline represents the median of the distribution. All of them are compared to the dose-response (DR) model. (c) represents the differences of RMSE between training and testing data for each model. 153
3.7 Modelled performance of a representative subject. Solid and dashed lines represent the variable dose-response model (DR) model and the two models offering the best generalisation. On this example, the training data set (80% of the data that combines training and validation subsets) and testing data set (20% of the data, the testing subset) areas are separated by the vertical solid line. Fitted parameters of the DR model were k_1 = -2.45e-05, $k3$ = -2.58e-09, $\tau_1 = 39$, $\tau_2 = 26$, $\tau_3 = 5$. Hyper-parameters of the PCR and ENET models were $n\ comp = 3$ and $\alpha = 0.28$, $\lambda = 0.02$. Note that in accordance with Elastic net (ENET) and principal component regression (PCR) models, only day of performances are displayed through DR. 157
3.8 Stacking ensemble learning using several fitness-fatigue and ML models. 164
3.9 Example of acceleration-velocity (A-V) profile modelled for a given player and a randomly selected game. Only plain dots were used for fitting the linear regression. 171
3.10 Evolution of A-V profiles fitted intercept and slopes over the period of study. Players are randomly selected. 172
3.11 Simplified diagram of (a) a recurrent neural network (RNN) cell and (b) a long short-term memory (LSTM) cell. 175
3.12 Distributions of averaged time-series forecasting performances. . . . 177
3.13 Example A-V profiles slopes forecasting using the uni-modal averaged ensemble. (a) represents the best prediction, (b) is the median prediction. 178
3.14 Distributions of model performances in a regression task. 181
3.15 Distributions of model performances in a classification task. 183

3.16 Distributions of model performances in a time-series forecasting using features extracted from raw data. 185
3.17 Distributions of model performances across models. 186

4.1 Basic framework for decision support system (DSS) building in sports. 191
4.2 maximal aerobic speed (MAS) test protocol on a 200m indoor track. The symbol a represents the start line, b is a photoelectric cell to reset the force platform records, both c are the two motion analysis sensor modules, d is the control panel and e are the cones laid every 20m and FP is the force platform recording area. 193
4.3 Mechanical stride changes during the MAS test. The top plots (a, b) represent changes in ground contact time (ms) (GCT) over speed and stride frequency respectively. The bottom plots (c, d) represent changes in leg stiffness ($kN{\cdot}m^{-1}$) (k_{leg}) over speed and stride frequency according to McMahon and Cheng 1990. In each figures, dots represent the group mean values, error bars the standard deviation in both x, y axes. The solid line is the regression line from the Bayesian linear model, surrounded by the 95% credible intervals. 199
4.4 $\dot{V}O_2$ – Stryd power output (PO) relationship during the incremental test. A strong and positive linear relationship was observed across participants. Lines represent each individual linear regression between $\dot{V}O_2$ and PO. 201
4.5 Comparison of PO estimated by the Stryd power meter and the force platform. The left plot (**a**) represents the strong positive relationship between the Stryd and the reference measures. The right plot (**b**) represents the averaged PO in response to speed, where the dotted line is the corrected Stryd PO (see text for details). 202
4.6 Bland–Altman plots for comparison of measurements between the force platforms (reference) and the power meter. Mean bias (middle dashed line), lower and upper limits of agreement (dashed lines) and their 95% confidence interval areas are represented. 206
4.7 Profile for a given runner. 208
4.8 Record power profile modelled according to Péronnet and Thibault 1989 for a given runner. In (a), black lines represent training sessions the power measured at each training session and the blue line is the record power profile estimated from time-trials. In (b), a heat map is displayed in order to represent distributions of the running power according to the time. 210
4.9 Daily performances of a given runner over the study period. 211
4.10 Summary of body mass measures with (a) average per position, (b) the progression of averaged mass of the loose-head prop position (c) a body mass monitoring of a given player playing at loose-head prop position. 213

4.11 Comparison of (a) hand grip and (b) counter-movement jump (CMJ) performances between players. 214
4.12 Example of training monitoring through some commercial global positioning system (GPS) features with (a) the evolution of acceleration-related features and (b) the evolution of velocity-related features over the last month. 215
4.13 Well-being dashboard with (a) a summary of well-being status for a given player, and (b) the evolution of each well-being item. 216

List of Tables

1.1 Original Borg rating of perceived exertion scale by Borg 1998 55
1.2 Original category-ratio scale of perceived exertion by Borg 1990 . . 56
1.3 Modification of the CR10 scale by Foster, Florhaug, Franklin, et al. 2001 57

2.1 Configuration of the three knee extension testing sessions 93
2.2 Parameters inference regarding changes in blood lactate concentrations in response to exercise. β^* represents standardised coefficient for each parameter of interest. 102
2.3 Parameters inference regarding changes in plasma cortisol concentrations in response to exercise. 104
2.4 Parameter inference regarding changes in averaged HR slopes $\bar{\alpha}^{HR}$ and averaged amplitude of recovery $\bar{HR}_{amp}$. 105
2.5 Parameters inference regarding rate of force development ($Nm.s^{-1}$) (RFD) and impulsion responses to exercise. 117
2.6 Parameters inference regarding the summated EMG signals (EMG_s) and the normalised averaged torque ($|\bar{T}|$) responses to exercise. Estimates β^* are standardised regression coefficients. 121
2.7 Parameters inference regarding MDF responses to exercise. 123
2.8 Parameters inference regarding slopes of MDF computed from VLat, vastus medialis (VMed) and rectus femoris (RFem) average (MDF_{quad}) at exercise. 124
2.9 Summary of the correlation analysis between four training loads (TL) quantification methods and the main physiological responses. . 134

3.1 Summary of independent variables. The two aggregation methods (impulse ans serial cumulative responses) are defined in Equations 3.1 and 3.2 148
3.2 Summary of models pairwise comparisons for generalisation and prediction abilities. β_{diff} represents the marginal mean difference of the RMSE distribution between the DR model and its comparison. 154

3.3 Summary of the predictive models. According to model families, criteria were averaged among folders and displayed with their standard deviation. For individual models, averaged values of hyper parameters are displayed along with lower and upper recorded values. The greatest performance among criteria is listed in bold type.* indicates the DR_I as the reference model and specification of its averaged parameters. 156
3.4 Summary of independent variables 169
3.5 Average mean absolute percentage error (MAPE) for each selected model. (*) additive seasonality, (**) multi-modal models required longer time-series. We limit the study of these models to time-series larger than 38 observations. 179
3.6 Summary of models performances for the classification task 183
3.7 Summary of models performances according to intercept and slope coefficients. $\overline{MAPE}$ represents the averaged MAPE over individuals and validation folders. The population represents either models computed over the group of players (G) or individually computed models (I). 187

4.1 Linear mixed modelling of $\dot{V}O_2$ and external mechanical power (W) ($\dot{W}_{ext}$) relationship. 198
4.2 Bayesian linear mixed models parameters estimates (see Appendix H.2 for details about models. 205

List of acronyms

$[cort_p]$
: plasma cortisol concentrations. 14, 95, 102, 103, 257

$[lact_b]$
: blood lactate concentrations. 38–41, 47, 95, 102, 106, 125, 126, 129, 134, 257

Δ [$\mathbf{0_2Hb}$]
: change in tissue oxyhemoglobin concentration. 96

Δ [$\mathbf{Hb_{diff}}$]
: oxygenation index, the difference between oxyhemoglobin and deoxyhemoglobin concentration. 14, 96, 98, 108–111

Δ [**HHb**]
: change in deoxyhemoglobin concentration. 96

Δ [**tHb**]
: change in total hemoglobin concentration, the sum of oxyhemoglobin and deoxyhemoglobin concentration. 14, 96, 110, 112

ANN
: artificial neural network. 84, 85

ASRM
: athlete self-report measures. 58, 59

ATP
: adenosine triphosphate. 42, 43, 50

A-V
: acceleration-velocity. 16, 170–173, 176–181, 183, 187, 188

bTRIMP
: Banister's training impulses. 38–40, 62, 126, 211

$C1$
: resistance exercise protocol: low intensity session. 14, 15, 92, 98, 102–104, 106, 108, 110, 115, 118, 120, 122, 123, 125, 129–131

$C2$
: resistance exercise protocol: moderate intensity session. 13–15, 92, 95, 98, 102–104, 106, 108, 110, 115, 118, 119, 122, 123, 125, 131

$C3$

resistance exercise protocol: heavy intensity session. 14, 15, 92, 98, 102–104, 106, 108, 110, 112, 115, 118, 122, 123, 125, 129–131

CK

creatine kinase. 48

Cm

mechanical cost of running ($J \cdot kg^{-1} \cdot m^{-1}$). 196, 197, 263, 264

CMJ

counter-movement jump. 18, 214

CO2

carbon dioxide. 94

CR10

0-10 category-ratio scale. 55, 56

CR100

0-100 "centimax" scale. 56

CV

cross-validation. 16, 139, 140, 143, 150, 151, 159, 164

DR

variable dose-response model. 16, 68, 69, 73, 74, 80, 144, 149, 152, 154, 157–160

DR_I

individually-computed variable dose-response model. 156

DSS

decision support system. 17, 34, 35, 191, 192

DV

dependent variable. 121

ECG

Electrocardiogram. 94

EE

energy expenditure (kcal). 42, 43, 52, 95, 106, 125, 129, 130, 134

ELISA

Enzyme linked immunosorbent assays. 95

EMG

Surface electromyography. 15, 19, 99, 118, 120–124, 257

ENET

Elastic net. 16, 144, 150, 152, 154–157, 159

$ENET_G$
group-computed elastic net regularisation. 152, 155, 156, 158

$ENET_I$
individually-computed elastic net regularisation. 152, 156, 158

EPL
English Premier-League. 32

EPOC
excess post-exercise oxygen consumption. 41–44

eTRIMP
Edward's training impulses. 39, 40, 45

F1-score
$F1 = 2\frac{\text{precision.recall}}{\text{precision+recall}}$. 182, 183

FFM
fitness-fatigue model. 13, 60, 61, 63–65, 67–69, 72, 73, 75, 77, 80, 83, 84, 154, 162, 163

FFMs
fitness-fatigue models. 60, 61, 64, 67–69, 73, 75, 76, 78–82, 84, 85, 138, 141, 144, 160–164

FFT
fast Fourier transform. 176, 184

FIT
flexible and inter-operable data transfer. 194

GCT
ground contact time (ms). 17, 192, 193, 195, 198–200, 203, 204, 207, 208, 263–265

GH
growth hormones. 47

GPS
global positioning system. 18, 32, 34, 52, 53, 166–168, 180, 183, 184, 187, 188, 191, 212, 215, 217

GRFs
ground reaction forces. 194, 195, 204

HI
High intensity. 92, 104, 112, 115, 123–125, 162

HR
Heart rate. 14, 19, 37–41, 44–46, 51, 54, 55, 94, 104, 105, 107, 125, 126, 208, 211

ICC
intraclass correlation coefficient. 197, 198, 203, 204, 263

i.i.d
independent and identically distributed. 141, 151

IL-6
interleukines-6. 48

impulsion
the area under the torque curve ($Nm.s$). 15, 19, 96, 112–115, 117, 127, 130, 133, 135, 257

IMU
inertial measurement units. 32, 50–53, 126, 166, 188

k_{leg}
leg stiffness ($kN{\cdot}m^{-1}$). 17, 196, 198, 199

KF
Kalman filter. 13, 73, 74

K-fold CV
K-fold cross-validation. 140, 141

LASSO
Least Absolute Shrinkage and Selection Operator. 150

LI
low intensity. 92, 104, 112, 124, 125, 162

LM
linear models. 82–84

LMM
linear mixed models. 82, 83, 102, 118, 123, 124, 133, 202, 203, 263, 265

LOOCV
leave-one-out cross-validation. 85, 139–141

LpOCV
leave-p-out cross-validation. 140, 141

LPS
local positioning systems. 52

LPT
linear position transducer. 50, 51, 126

LQ
linear quadratic. 161

LSS
leg spring stiffness. 192, 193, 196, 207, 208, 264

LSTM
long short-term memory. 16, 173, 175–177, 179, 182, 184, 187

lSVC
linear support vector classification. 173, 182, 187

LT
lactate threshold. 37, 40

luTRIMP
Lucia's training impulses. 40, 41, 45

L-V
load-velocity. 51

$m\dot{V}O_2$
muscular oxygen uptake ($ml.min^{-1}$). 96, 108, 125

MAE
mean absolute error. 63, 64, 139

MAPE
mean absolute percentage error. 20, 63, 176, 179, 184, 187, 188

MAS
maximal aerobic speed. 17, 52, 193, 198, 200, 202, 264, 265

MDF
median frequency. 15, 19, 99, 122–124, 129, 134, 258

MDF_{quad}
MDF computed from VLat, VMed and RFem average. 19, 122–124

ME
mechanical efficiency (%). 196, 198, 203

MI
moderate intensity. 92, 104, 112, 115, 118, 124

ML
machine-learning. 33, 34, 84–86, 138, 141, 144, 160, 163, 164

MUs
motor units. 120

MVC
maximum voluntary contraction. 93, 120, 125, 127

NIRS
near-infrared spectroscopy. 96

O2
oxygen. 37, 94

OOS
out-of-sample. 141, 142

PC
principal components. 128, 133, 150

PCA
principal component analysis. 83, 128, 129, 133, 135, 136, 150

PCR
principal component regression. 16, 144, 150, 152, 154, 155, 157, 159

PCr
phosphocreatine. 42, 43, 50, 96

PCR_G
group-computed principal component regression. 152, 156–158

PCR_I
indivual-computed principal component regression. 152, 156

PerPot
Performance Potential. 13, 75–77

PO
power output. 17, 44, 45, 192–194, 200–204, 207, 208, 264, 265

POMS
Profile of Mood States. 58

PSD
power spectral density. 99

R^2
coefficient of determination. 156–158

RE
resistance exercise. 43, 48–51, 89, 94, 103, 104, 118, 125, 126, 129, 133–136

REc
running economy. 207

RESTQ
Recovery-Stress Questionnaire Athletes. 58

RF
random forest. 144, 151, 156

RF_G
group-computed random forest model. 155, 156, 158

RF_I
indivual-computed random forest model. 155

RFD
rate of force development ($Nm.s^{-1}$). 19, 96, 112, 115, 117, 118, 124, 126, 127, 129, 135, 136

RFD_{0-100}
rate of force development from the onset to 100ms of exercise ($Nm.s^{-1}$). 15, 96, 112, 114–118, 125, 134, 257

RFD_{peak}
maximal rate of force development computed from 20ms time intervals($Nm.s^{-1}$). 15, 96, 112, 114–118, 125–128, 133, 134, 257

RFem
rectus femoris. 19, 25, 99, 118, 122, 123

RM
repetition maximum. 13, 49–51, 92, 93, 101, 125, 134, 135

RMS
root mean square. 99, 118–120

RMSE
root mean square error. 16, 19, 63, 139, 152–157

RNN
recurrent neural network. 16, 173, 175–177

RPE
rate of perceived exertion. 54–57, 88, 89, 126, 130, 133, 146

RPP
record power profile. 209, 211

RSS
residual sum of squares. 65, 77, 182

RT
resistance training. 46–48, 50, 51, 88, 126, 135, 191, 217

sRPE
session rate of perceived exertion. 56–58, 88

SRS
spatially resolved spectroscopy. 96

STFT
short-time Fourier transformations. 14, 99, 100, 123

TL
training loads. 19, 31, 33, 36, 37, 39, 41, 43–46, 48, 50–54, 56, 57, 59, 62, 88, 89, 126–130, 133–136, 145, 160–163, 166, 190, 217

TSI
tissue saturation index (percentage of ratios of absorbance). 14, 96, 108–111, 125

TSIdx
training stress index. 89

TSS
training Stress Score. 44, 45

TUT
time under tension. 50, 101, 106, 112, 113, 123, 125, 126, 134

T-V
torque-velocity. 13, 89–92, 101, 103, 125, 134–136

VARIMA
vector autoregressive moving average. 176

VBT
velocity-based training. 51, 135

VL
volume load. 48–50, 89, 106, 125, 126, 129, 130, 133, 135

VLat
vastus lateralis. 14, 15, 19, 25, 96, 99, 100, 108, 118, 119, 122, 123

VMed
vastus medialis. 19, 25, 99, 118, 122, 123

$\dot{V}O_2$
oxygen uptake ($ml.min^{-1}$). 13, 14, 17, 20, 37, 38, 40–42, 44, 94, 95, 105–107, 125, 196–198, 200, 201

$\dot{V}O_2max$
maximal oxygen uptake ($ml.min^{-1}$). 37, 38, 42, 192, 193

$\dot{W}_{ext}$
external mechanical power (W). 20, 192, 196–198, 202, 207, 208, 211

W_{mech}
: mechanical work (J). 15, 112, 114

$\dot{W}_{met}$, **W**
: Metabolic power (W). 196, 208

Glossary

athletic performance
A performance conditioned by physical, technical, physiological, psychological and cognitive skills under the influence of its environment. 31–36, 41, 59, 60, 64, 65, 73, 75, 77, 78, 80–86, 88, 135, 138, 141, 144, 159, 160, 162–164, 166, 167, 170, 176, 179, 188–192

endurance sports
Any sport in which there is a requirement to sustain an activity level while enduring a level of physical stress. 38, 46, 161

features
A feature is a measurable property of an object. In our case, it relates to an exercise-related variable. 32

generalisation
The ability of a trained model to accurately predict on examples that were not used for training. 138, 191, 217

multivariate
A model or analysis that considers at least two explanatory variables. 34, 163, 179, 180, 183, 184, 187, 217

overfitting
An over-trained model, which tends to memorise each particular observation thus leading to high error rates when predicting on unknown data. 138

resistance training
Any exercise that causes the muscles to contract against an external resistance with the expectation of increases in strength, power, hypertrophy, and/or endurance. 11, 35, 36, 39, 46, 47, 56, 87, 89, 126, 162, 188

RR
Intervals between successive heartbeats. 94

underfitting
An inflexible model unable of capturing noteworthy regularities in a set of exemplary observations. 138

Introduction

Relationships between training effects on athletic performance are a great challenge that coaches, sports scientists, and professionals supervising athletes focus on. Since the first competitions, athletes trained themselves according to training plans, sequenced in various development cycles. Training programming came thus with objectives defined by athletes along with coaches and medical staff. It is of importance that athletes should be fully involved in their preparation. One can say that they may consider training programs as a lifeline on which they lay back on, leaving their fate to the hand's coach and doing their utmost to reach the planned goals. That mutual commitment brings a great responsibility for prescribers (coaches) shoulders when the pursuit of the highest performance is engaged.

In athlete monitoring, individual skills are assessed through various tests, performed either in laboratory or ecological conditions. Evaluations highlight the strengths and weaknesses of each athlete, constituting a basis for any individual training programming. Then, a close follow-up of athletes progression over seasons comes with regular and repeated evaluations, as a requisite for any decent athlete status monitoring. In order to map the effects of training and the performance outcomes, we define a TL index that represents a quantitative measure of the effort done by the athlete and the underpinning induced stress. Accordingly, TL are commonly dissociated into i) an *external* load defined by the work completed by the athlete, independently of his internal characteristics (Wallace, Slattery, and Coutts 2009), and ii) an *internal* load that corresponds to the psycho-physiological stresses imposed on the athlete in response to the *external* load (Impellizzeri, Rampinini, and Marcora 2005). In other words, TL being quantified in both quantitative and qualitative ways, it would allow answering the questions: what amount of effort has been done by the athlete, and what does it really means, among others, in terms of physiological adaptations? These central questions will be answered all through the manuscript. Before that, let us define some terms. The term *loads* has a mechanical, physical meaning and describes a force. Strictly speaking, it should be accompanied with the SI-derived unit of the newton (N) and not be used for describing any training-related variables, which more or less mechanical meaning (Staunton, Abt, Weaving, et al. 2021). Conversely, a *stress* might be considered as state variation in human functions induced by exercise. It is likely more generic and free of any specific SI unit. However, for conformity with the literature, we will consistently use the terms of *training loads* but with a *stress* connotation behind them.

Assessing an athlete performance often requires particular sessions and is, therefore, time-consuming. While time may be in-expandable regarding the preparation for a competition schedule, coaches may benefit from training observations to evaluate individual progression. In this way, the rise of wearable sensors for measuring accurately the least effort performed by athletes has greatly facilitated and encouraged athlete status monitoring. Intended to measure an overall activity during training sessions and competitions, they provide useful insights that may be used for both TL and performance quantification in any sports and ecological conditions. Training programs thus become evolutive and daily adjustable in order to be optimal for each athlete. Yet, wearable sensors such as GPS and inertial measurement units (IMU) require practitioners to deal with a large amount of data and to fully understand what and how are the variables measured. Thus, it might imply a change in usages of coaching, becoming more data-driven or data-informed and involves close collaborations between training, medical staff and data experts.

Beyond the singular satisfaction brought by athletic performance to athletes, significant economic issues arise therefrom. On one side, winning world-class competitions such as world championships and Olympic games raise the sport's development within the country, reinforcing national sport organisations while at the same time improving the nation credibility at an international level. On the other side, injuries are part of sports performance. In economic terms, injuries represent major losses for clubs and sports organisations. As an example, *Forbes* estimated the cost of injuries in English Premier-League (EPL) of about 267 millions dollars (McMahon 2019). By considering wage bills and prize money, a recent study from Eliakim, Morgulev, Lidor, et al. 2020 estimated the financial damages caused by injuries in EPL only, up to £45 million sterling.

Fortunately, the training effects of athletic performance are a major topic in exercise physiology researches. Largely studied for years, the theoretical physiological mechanisms related to exercise allows us to understand and estimate what physiological adaptations are susceptible to occur following a training session and their aftereffects on athletic performance. Scientists attempted to model the effects of training on physical performance on a physiological basis, initially using system model frameworks (Banister, Calvert, Savage, et al. 1975; Calvert, Banister, Savage, et al. 1976). Sometimes named "biocybernetics" models, they aim at describing and predicting performance outcomes using states features, built from more or less elaborated functions intended to represent some basics of biological processes (*e.g.* super-compensation process).

However, human is made of complex biological systems acting as a network in which several processes exchanges between each other at various orders (Bazyler, Abbott, Bellon, et al. 2015; Lambert, Gibson, and Noakes 2005). Not surprisingly, traditional system models of training effects lack of descriptive and predictive powers (Hellard, Avalos, Lacoste, et al. 2006), since they resume training effects to a very few features if it is not a single one in most cases. That makes such models

useful for modelling performance trends from TL dynamics but not valuable for predicting athletic performances with accuracy and finding an optimal training sequence accordingly.

Beyond models used for athletic performance purposes, information forms the basis of any modelling process. In sports, information stems from various sources, being objectively or subjectively measured. While TL are training-related parameters at the basis of former system models, any other parameters related to training (e.g. technical, environmental, social, psychological, nutritional) may bring valuable insights in the modelling under the multifaceted of athletic performance. Hence, all available information for understanding the relationship between training and performance should be considered in modelling processes. However, objective *external* TL measures are still often used in isolation without any consideration for other aforementioned information. This might be a significant inherent limitation of system models used so far, whereas many statistical approaches may benefit from the richness of available data.

Statistics and computer science show a great attraction over the last two decades in sports science and sports analytics, with high predictive power in particular for solving complex non-linear problems (Carrard, Kloucek, and Gojanovic 2020; Edelmann-Nusser, Hohmann, and Henneberg 2002; Mitchell, Rattray, Fowlie, et al. 2020). Machine-learning modelling approaches seek to approximate a function that maps the input (*e.g.* a combination of predictors) to the output (*i.e.* an athletic performance), lowering errors between predictions and observations in their way. Choosing a model (or a class of models) of interest for solving a particular problem will depend on the structure, properties of the data such as the sample size and the presence of temporal dependencies between observations.

Despite being extensively used in several domains, one should pay attention to potential drawbacks behind some machine-learning (ML) approaches. On one side, system models come with a strong hypothesis and are highly interpretable models for the benefit of coaches. On the other side, hypothesis-free ML models might suffer from direct such interpretation, particularly the so-called *black-box* models. To draw an interpretation of model parameters remains essential for practitioners, aiming at understanding athlete responses to exercise for the optimisation of training programs. Yet, it does mean that ML should be discredited but rather used consistently. Also, since system models rely on a hypothesis based on macro-biological principles, they might be integrated with machine-learning algorithms for the benefit of bio-physiological expertise along with powerful modelling frameworks.

In summary, sports-related issues are sufficient to consider the relationships between the effects of training and athletic performances as of major interest, sitting as a valuable research topic that has already been studied for years but has not been elucidated so far.

Understanding relationships between training and athletic performance implies:

1. Having rigorously collected data that include all relevant information for

athletic performance comprehension.

2. Performing a robust and valuable modelling approach for solving complex problems in which key variables are identified for explaining training outcomes.

Furthermore, in a dose-response framework, training effects were considered through direct and cumulative principles while considering dissipating effects over time. However, little is known about delayed training effects, despite the non-linear aspects of physiological adaptations to training are stated in the literature. Since system models benefit from control theory, extending the former transfer functions to more complex functions might better represent physiological adaptations to exercise and, therefore, athletic performance outcomes.

While the former system models are used for description and prediction purposes, their ability to predict future performances is rarely robustly assessed. It questions the pertinence of drawing interpretation from models, which might be partly flawed and intended for optimising training programs.

Finally, companies that develop sports wearable sensors such as GPS provide summarised features of measured data, either based on scientific or practical evidence. It means that customers are restricted to the company's choice in the feature provided, whereas new meaningful features might be extracted from the raw data using signal processing techniques.

The work presented in this manuscript is based on two concurrent objectives. First, it advances athletic performance modelling by combining exercise physiology knowledge from theory to practice and using appropriate statistical methods to solve complex issues. The latter relates to developing a DSS intended for elite athletes, performance and medical staff.

Therefore, the manuscript reviewed the methods used for athletic performance modelling and provided more appropriate modelling approaches from statistics and computer science. In methodological terms, careful attention is paid to the selection of optimal models (*i.e.* models that efficiently perform according to a given context) and their generalisation capability. From an applied perspective, the method fosters statistical and ML approaches based on training parameters aggregates for making predictions of athletic performance in ecological conditions.

Since many models may provide similar results for athletic performances prediction purposes (Wolpert and Macready 1997), we provided some robust comparisons using field data and highlighted both limits of each approach and appropriate solutions to the problem.

Accordingly, it is hypothesised that multivariate approaches using features from various sources are better suitable than former system models for predicting athletic performances. In addition, extracting new information from GPS raw data might be a valuable alternative, making the most of wearable technologies that are prone to benefit from future developments.

This manuscript consists of four chapters, each of them addressing particular issues. The first chapter provides a state-of-the-art of contemporary methods for quantifying training loads and a narrative review of the models used for athletic

performance modelling. The second chapter is intended to advance the training load quantification methods applied to resistance training by conducting investigations in laboratory conditions. The third chapter is directly related to modelling applied to athletic performance in elite sports. Various applications provide key insights and guide further applications in sports. The last chapter provides some DSS intended for athletes, coaches and sports structures. The DSS includes data-visualisation, team management and predictive analysis tools applied to various populations and sports.

1. State of the art

Table of contents

1.1 Training load, definition and quantification 36
1.1.1 Objective measures 37
1.1.1.1 Quantification in endurance sports 37
1.1.1.2 Quantification in resistance training 46
1.1.1.3 Quantification in field sports: contribution of micro-technology 51
1.1.2 Subjective measures 54
1.1.2.1 Rate of perceived exertion 54
1.1.2.2 Athlete self-report measures 58
1.2 Modelling the effects of training 59
1.2.1 A collection of Fitness-Fatigue models 59
1.2.1.1 Original fitness-fatigue model: a two-components model 61
1.2.1.2 A linear time-varying Fitness-Fatigue model 64
1.2.1.3 A preload concept for improving Fitness-fatigue models 65
1.2.1.4 A variable dose-response model 67
1.2.1.5 Fitness-fatigue and delayed effects 69
1.2.1.6 Improving fitness-fatigue models with Kalman filter 73
1.2.2 Analogies with the fitness-fatigue framework 75
1.2.2.1 Training effects and performance potential 75
1.2.2.2 Indirect pharmacodynamic models for athletic performance 78
1.2.3 Contributions from statistics and computer science 81
1.2.3.1 Statistics and multivariate linear models 81
1.2.3.2 A machine-learning approach to the problem 84
1.3 Conclusion 86

1.1. Training load, definition and quantification

Training load has become a gold standard measure for representing how much stress is induced by physical exercise. Essential in any monitoring process, TL stems from various sources of data, and it can be assessed in several ways, according

to exercise specific properties. Objective and subjective estimates of TL are the two main approaches used by the community.

In the following sub-sections, we present the main objective and subjective measures used for quantifying training loads. For each of them, an underpinning physiological rationale is presented upstream.

1.1.1. Objective measures

1.1.1.1. Quantification in endurance sports

Physiology-based measurements

Heart rate kinetics during and after exercise have been extensively studied so far (Bunc, Heller, and Leso 1988; Cooper, Berry, Lamarra, et al. 1985; Engelen, Porszasz, Riley, et al. 1996; Karvonen and Vuorimaa 1988; Knuttgen, Petersen, and Klausen 1971; Schneider, Wing, and Morris 2002). Since HR correlates with $\dot{V}O_2$ at several sub-maximal exercise intensities (Cooper, Berry, Lamarra, et al. 1985; Maritz, Morrison, Peter, et al. 1961; Verma, Sidhu, and Kansal 1979), it has become one of the most used physiological parameters for exercise intensity prescription and monitoring purposes (Achten and Jeukendrup 2003; Borresen and Lambert 2009). HR monitoring would thus offer an objective, continuous and non-invasive measure of exercise intensity through acute cardiopulmonary responses (Borresen and Lambert 2009).

While a simplified relationship between $\dot{V}O_2$ (or indirectly HR) and work rate is assumed where $\dot{V}O_2$ increases as a linear function of work rate, the same cannot be said for any exercises performed over lactate threshold (LT) or exercises that involve a sustained lactic acidosis (Barstow, Casaburi, and Wasserman 1993; Gaesser and Poole 1996). In those cases and according to the energetic pathways committed, a slow component of $\dot{V}O_2$ (Henson, Poole, and Whipp 1989; Jones and Poole 2005) takes place in order to supply the energy demand; leading to an increased oxygen (O2) cost per watt expected based on of a linear relationship between work rate and $\dot{V}O_2$ during an incremental exercise test. The magnitude of that slow component increases along with the intensity above LT until reaching maximal oxygen uptake ($ml.min^{-1}$) ($\dot{V}O_2max$) or exhaustion (Gaesser and Poole 1996). Hence, for a constant load exercise, a break-point in the linear relationship occurs at LT that marks the onset of the slow component phase (Barstow and Molé 1991; Henson, Poole, and Whipp 1989; Roston, Whipp, Davis, et al. 1987; Zoladz, Duda, and Majerczak 1998) identified by an exceeding $\dot{V}O_2$. That is obviously of importance when considering exercise intensity prescription through absolute or relative $\dot{V}O_2$ levels. As regards HR kinetics at exercise, an expected slow increase in HR is also observed during constant work-rate exercise (Zuccarelli, Porcelli, Rasica, et al. 2018). Yet, HR time course and magnitude differ from those of $\dot{V}O_2$, due to different mechanistic determinants behind HR and $\dot{V}O_2$ regulation processes (Zuccarelli, Porcelli, Rasica, et al. 2018).

Typical signatures of HR kinetics are commonly explained according to individual training states. For instance, trained athletes with lower resting HR value than untrained counterparts express faster HR kinetics. Similarly, a greater $\dot{V}O_2max$ induces a shorter time course and adjustment for both $\dot{V}O_2$ and HR kinetics Bunc, Heller, and Leso 1988; Hickson, Bomze, and Hollozy 1978 at exercise. On this basis, using HR as a reflection of $\dot{V}O_2$ and thus exercise intensity has been employed as a valuable parameter in prescribing sub-maximal exercises performed at constant work rates (Achten and Jeukendrup 2003; Borresen and Lambert 2009; Karvonen and Vuorimaa 1988) such as performed in the majority of endurance sports. For programming purposes and to carry analyses out through athletes comparisons, a relative measure of HR is usually preferred to absolute measure. A percentage of maximal HR ($\%HR_{max}$) is a common method for calculating exercise intensities of athletes (see Equation 1.1) (Karvonen and Vuorimaa 1988). Since the resting heart rate increases with age whereas the maximal heart rate decreases, $\%HR_{max}$ is usually the most appropriate method for exercise prescription, preferred to relative HR to peak value or HR according to a percentage of maximum METs (Karvonen and Vuorimaa 1988). Notwithstanding this, we can estimate the exercise intensity of an endurance training session such as

$$\%HR_{max} = \frac{HR_{work} - HR_{rest}}{HR_{max} - HR_{rest}} \,. \tag{1.1}$$

For monitoring purposes, training load indexes are usually calculated from the product of exercise intensity and training volume. While the calculation of training volume is quite simple, relying on total distance covered or time spent at exercise, the determination of exercise intensity remains much difficult. As presented above, the relationship between work rate and the resulting metabolic stress is non-linearly related in several cases. This non-linearity may be illustrated by an exponential increase of blood lactate concentrations ($[lact_b]$) as a function of work rate and $\dot{V}O_2$ (Davis, Rozenek, DeCicco, et al. 2007; Farrell, Wilmore, Coyle, et al. 1979; Freund, Oyono-Enguelle, Heitz, et al. 1986; Hughson, Weisiger, and Swanson 1987; Hurley, Hagberg, Allen, et al. 1984; Wasserman, Beaver, and Whipp 1986). Therefore, quantifying training loads objectively while depicting the physiological adaptations to exercises should consider these relationships but remains challenging.

One of the first and most famous system of training quantification relying on HR comes from the work of Banister and Hamilton 1985, the so-called "training impulse" (Banister's training impulses (bTRIMP)). Similarly to (Karvonen and Vuorimaa 1988) using $\%HR_{max}$ for exercise prescription, bTRIMP relies on the average fractional elevation of the maximum HR range for a measure of intensity denoted HR_r, multiplied by the exercise duration and weighted by a non-linear coefficient k that accounts for the greater taxes and adaptations involved by high

exercise intensities (see Equation 1.2). Training impulse is defined as following:

$$\begin{aligned} bTRIMP &= D\ HR_r\ k, \\ with\ HR_r &= \frac{\left(\frac{1}{N}\sum_{n=1}^{N} HR_{ex} - HR_{rest}\right)}{HR_{max} - HR_{rest}}\ . \end{aligned} \tag{1.2}$$

Here, $k = 0.64e^{1.92\ HR_r}$ for men, and $k = 0.86e^{1.67\ HR_r}$ for women. The parameters D and HR_r denote the duration of exercise and the average fractional elevation of the maximum HR range, respectively. The parameter k positively weights the effort at high HR in line with the exponential increase in $[lact_b]$. HR_{ex} is the HR measured at exercise. Both amplitude and time constant of HR kinetics vary according to gender (Green, Hughson, Orr, et al. 1983).

From this basis, bTRIMP allows for quantify training in continuous situations (*i.e.* where intensity of exercise remains quite steady or where HR changes slowly operates through the exercise / training session). It is also possible to quantify the training for particular phases of the training session, if intensity is held long enough to observe valuable changes in HR. In this case, the overall TL would be given by the sum of each bTRIMP phase scores (Banister and Hamilton 1985; Garcı+a-Ramos, Feriche, Calderón, et al. 2015).

The bTRIMP method for quantifying TL has several limitations. First, it is not a valuable estimate of TL neither for intermittent exercises since exercise and rest phases cannot be discriminated if we refer to Equation 1.2, nor for resistance training exercises since HR does not increases proportionately with load, that is resistance exercise intensity (Borresen and Lambert 2009). In addition, the use of generic parameters in the formulation of the weighting coefficient k neglects the individual differences in $[lact_b]$ responses to exercise, in particular when training levels and $[lact_b]$ tolerance greatly vary between athletes.

Alternatives to the original training impulse quantification from Banister and Hamilton 1985 have been proposed. Instead of weighting high intensities using a coefficient based on the kinetics of one physiological parameter (i.e. such as given by Equation 1.2), Edwards 1993 chose to split the Edward's training impulses (eTRIMP) span into five equated zones. An exercise training impulse eTRIMP is calculated by multiplying the accumulated duration in each zone of intensity to its corresponding weighting factor K_e, defined in the sequel:

$$K_e = \begin{cases} 1 & \text{if} \quad \frac{\overline{HR}_{ex}}{HR_{max}} \in [0.5, 0.59] \\ 2 & \text{if} \quad \frac{\overline{HR}_{ex}}{HR_{max}} \in [0.6, 0.69] \\ 3 & \text{if} \quad \frac{\overline{HR}_{ex}}{HR_{max}} \in [0.7, 0.79] \\ 4 & \text{if} \quad \frac{\overline{HR}_{ex}}{HR_{max}} \in [0.8, 0.89] \\ 5 & \text{if} \quad \frac{\overline{HR}_{ex}}{HR_{max}} \in [0.9, 1] \, . \end{cases} \tag{1.3}$$

Here, K_e denotes the score attributed to each HR zone, $\overline{HR}_{ex}$ denotes the averaged HR over the phase of interest and HR_{max} is the individual maximum HR. This method came with the need of an objective estimate of exercise intensity and thereafter, an objective estimate of training load for intermittent exercises. While bTRIMP is not suitable for intermittent exercises, eTRIMP would be a more appropriate method that discriminates exercise according to the aforementioned zones of intensity (see Equation 1.3). However, it is of importance to note that eTRIMP weights TL according to intensity zones in a linear way, thus neglecting the non-linear reflect of physiological responses (HR and $\dot{V}O_2$) to exercise above LT (Borresen and Lambert 2009). Yet, an extension of eTRIMP that includes an exponential weighting factor has been proposed by Stagno, Thatcher, and Van Someren 2007.

A similar but somewhat simpler has been suggested by Lucia, Hoyos, Carvajal, et al. 1999 in cycling. The author established three fixed exercise zones based on HR equivalences to either lactate thresholds (see Equation 1.4, $[lact_b]$ values being measured in $mmol.L^{-1}$) (Lucia, Hoyos, Carvajal, et al. 1999), or ventilatory thresholds (Lucia, Hoyos, Santalla, et al. 2003), respectively. Again, multiplying the time spent in each training zone to the corresponding weighting coefficient K_{lu} and then summed for each training bouts gives the total Lucia's training impulses (luTRIMP) score. Compared to eTRIMP method, figuring the exercise intensity out according to only three zones do not allow for a precise representation of the real exercise intensity performed (e.g. a same value is given for both LT and maximal aerobic power or velocity exercise intensities). According to $[lact_b]$ or HR equivalences, the value taken by K_{lu} also linearly weights the TL index without considering the exponential nature of the exercise demand at high intensities.

$$K_{lu} = \begin{cases} 1 & \text{if} \quad [lact_b] < 2 \\ 2 & \text{if} \quad 2 \leq [lact_b] < 4 \\ 3 & \text{if} \quad [lact_b] \geq 4. \end{cases} \tag{1.4}$$

Quantification methods of TL based on training impulses come with their respective limits. First, they are dedicated to exercises performed at steady-state and sub-maximal intensities, and therefore they do not seem to be valuable for intermittent exercise training (Tschakert and Hofmann 2013). Beyond the nature of exercise (i.e. continuous or intermittent), none of the aforementioned methods is expressed as a measure of density that accounts for passive rest or pause time during sessions. In addition, because they are HR-based methods, a HR monitor is required whatever the activity. For some sports such as swimming, wearing HR monitors may be detrimental to the athletic performance, thus questioning such methods for training load quantification purposes.

While $[lact_b]$ let the computation of luTRIMP scores through HR equivalences, some authors opted for sport-specific equivalences (Hellard, Scordia, Avalos, et al. 2017; Mujika, Busso, Lacoste, et al. 1996; Thomas, Mujika, and Busso 2008). Initiated by Mujika, Busso, Lacoste, et al. 1996, authors asked athletes to perform a few swimming incremental tests to exhaustion through the season. $[lact_b]$ were measured all along the tests to determine five swimming velocity zones. Then, a weighting factor K_m is attributed individually to each zone, such as :

$$K_m = \begin{cases} 1 & \text{if} \quad bLact < 2 \\ 2 & \text{if} \quad 2 \leq bLact < 4 \\ 3 & \text{if} \quad 4 \leq bLact < 6 \\ 4 & \text{if} \quad bLact \geq 6 \\ 5 & \text{if velocity is maximal} \end{cases}$$

Then, the session TL (expressed in arbitrary units) is given by the sum of distances (in kilometres) swam at each training intensity, weighted by their respective factor K_m. This method allows for weighting TL estimates according to a physiological basis, even though it is in a linear way. Since this method allows for estimating TL using a simple chronometer and a few testing sessions for the calibration of individual intensity zones, it might be applicable in most conditions.

While all of the aforementioned methods quantify TL from exercise parameters, excess post-exercise oxygen consumption (EPOC) has also proved its worth for estimating "fatigue" induced by exercise and has been commercialised so far (Rusko, Pulkkinen, Saalasti, et al. 2003). As its name suggests, EPOC comes from an excess of $\dot{V}O_2$ after exercise, marked by elevated levels above resting values for some period of time. The concept behind EPOC arises from the very first statement of Hill and Lupton 1923 in which a post exercise oxygen debt exists in order to repay the oxygen deficit incurred early at exercise onset, and ascribed to the oxidative removal of lactate. A little later, Margaria, Edwards, and Dill 1933 attributed the

oxygen debt to the lactacid component induced by the glycogen synthesis from lactate and an alactacid caused by other factors. Finally, Gaesser and Brooks 1984 comes with a more realistic causality with EPOC, which is not only the short component of $\dot{V}O_2$ increases but also a prolonged one that may persists for up to 24 hours (see Figure 1.1).

In their review article, Børsheim and Bahr 2003 reported that the magnitude of EPOC was curvilinearly related to the intensity of exercise, with a break point around intensity corresponding to 50–60% of $\dot{V}O_2$max during constant work rate exercises. On the other side, the relationship between the magnitude of EPOC and the exercise duration seems to be mostly linear Børsheim and Bahr 2003 all along the time course. Nevertheless, individual differences in EPOC responses to a same relative exercise stimulus exist Børsheim and Bahr 2003 according to the heterogeneity of responsiveness to exercise in healthy people (Bouchard and Rankinen 2001).

Before addressing the effects of different types of training on EPOC, let us lay the foundations of some mechanisms underpinning EPOC. First, main mechanisms responsible for the short component identified so far come from a global myocellular homeostasis that includes replenishment of oxygen stores, adenosine triphosphate (ATP) and phosphocreatine (PCr) resynthesis, lactate oxidation and removal, restoration of fluid balance and fuel stores, increased body temperature (Børsheim and Bahr 2003). A note of importance concerns the restoration of fuel stores according to an increase of lipid oxidation for "saving" carbohydrate energy sources during recovery (Kiens and Richter 1998). That shift in substrate utilisation justifies a high priority given to muscle glycogen resynthesis, while the lipid oxidation from intramuscular triglycerides and free fatty acid is attributed to the restoration of full requirements (Egan and Zierath 2013). This mostly contributes to the prolonged component of EPOC, according to plasma concentrations of catecholamines that increase during exercise and which regulates the triglyceride/fatty acid cycle (Børsheim and Bahr 2003).

According to the literature, we can argue that both exercise intensity and duration impact EPOC. However, for a similar intensity and duration, differences in EPOC can be observed according to the type of training (*e.g.* continuous or intermittent training and the modality of muscle contraction). Substantial increases of EPOC after intermittent exercise sessions were reported by authors (Almuzaini, Potteiger, and Green 1998; Kaminsky, Padjen, and LaHam-Saeger 1990). In addition, when comparing continuous and interval exercises at similar controlled energy expenditure (kcal) (EE), magnitude and duration of EPOC were significantly greater in interval training sessions than in continuous ones (Cunha, Midgley, McNaughton, et al. 2016; Jung, Hwang, Kim, et al. 2019). These results support the hypothesis that EPOC is greatly dependent on exercise intensity. According to the mechanisms responsible for EPOC aforementioned, we can logically deduct that increase in EPOC due to higher exercise intensities may be explained by the energy cost to resynthesise glycogen from lactate, an increase in core temperature due to a metabolic activity,

a resynthesis of ATP and PCr stores as well as inflammatory responses through changes in cytokine release (Børsheim and Bahr 2003). Consequently, a greater exercise intensity may greater affect the homeostasis state, reflected by a larger EPOC. This result is manifestly of importance when aiming to estimate the effects of exercise on the body. Off-exercise periods should thus be considered in any monitoring process for a better estimation of the training effects.

Some authors have compared EPOC in response to aerobic and resistance exercises, according to an equated EE (usually estimated through indirect calorimetry such as gas exchange analysis) across exercises. Greater EPOC were observed following intermittent resistance exercise (RE) and high-intensity intermittent exercises than continuous aerobic exercises (Burleson, O'Bryant, Stone, et al. 1998; Gillette, Bullough, and Melby 1994; Greer, Sirithienthad, Moffatt, et al. 2015). However, a precise estimate of EE during both resistance and continuous aerobic exercises remains challenging and should be considered when interpreting EPOC kinetics in these conditions.

A simpler method for equating the total work across RE is based on the product of relative intensity and volume of RE (this method will be further detailed in this chapter). It eases the balancing of total work in RE in ecological condition without necessitating the measures of the metabolic pathways. In this way, the results reported by Thornton and Potteiger 2002 agree with the literature, namely a significant effect of exercise intensity on the magnitude of EPOC though a similar EE.

In brief, it is clear that EPOC is impacted by exercise duration, intensity and modality of muscle contraction. Hence, it may be a valuable parameter for estimating stress induced by exercise in TL monitoring purposes. In addition, sex and training status may impact EPOC. Menstrual cycles should thus be considered when controlling EE in women, and absolute measures of intensity should be avoided in favour of relative measures of intensity (Børsheim and Bahr 2003).

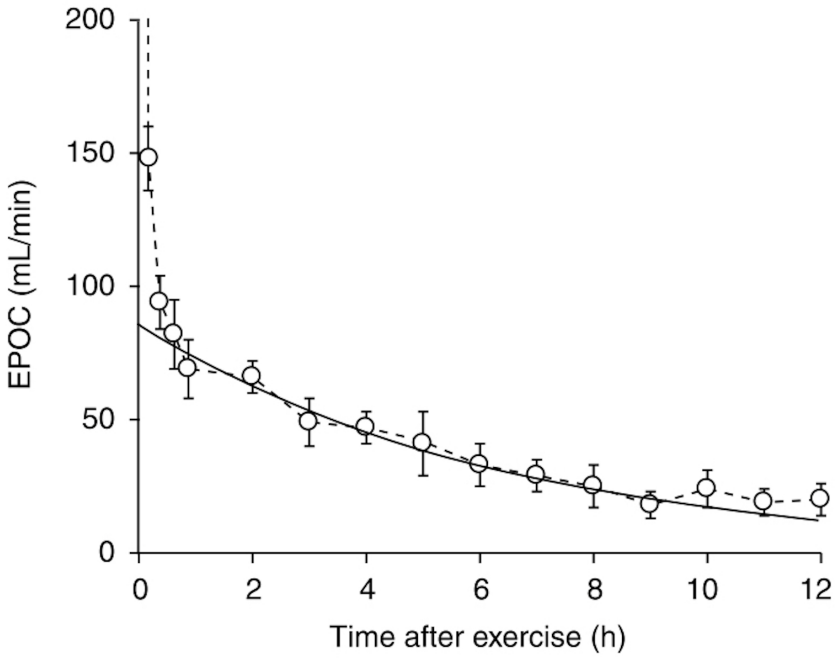


Figure 1.1. – Excess post-exercise oxygen consumption (EPOC) after exhaustive submaximal exercise (71-80 minutes at 69-78% of $\dot{V}O_2max$), taken from Børsheim and Bahr 2003. Circles represent the recorded $\dot{V}O_2$ values and the solid line shows the prolonged component.

Take-home message Heart rate is a key parameter for estimating exercise intensity and therefore training loads. Numerous drawbacks behind the use of HR still have to be taken in consideration, limiting its usefulness for supra-maximal and intermittent exercises. In addition, environmental (*e.g.* temperature, humidity), psychological and physiological factors are prone to affect the HR – exercise intensity relationship and should also be considered for longitudinal analysis purposes (Lambert, Mbambo, and Gibson 1998). But be it used through TRIMP calculations or EPOC estimates according to the $\dot{V}O_2$ – HR relationship, HR allows for an objective estimate of training loads according to physiological changes occurring during exercise.

External parameter-based measurements.

By analogy to Banister's training impulses, training Stress Score (TSS) has been proposed by Coggan 2003 as an *external* TL indicator. First developed in cycling, TSS is based on PO measurements and has become a gold standard for quantifying TL in cycling and more recently in running (Van Dijk and Van Megen

2017). Before going deeper with TSS, let us explain why it is based on PO and why PO taken alone is of limited significance for quantifying TL.

At first glance, instantaneous PO seems to be a valuable measure of exercise intensity, not impacted neither by elevation, nor by surface quality or any external factors that can influence velocity. Despite it would logically be of direct interest within a training monitoring process, it requires practitioner to deal with a large amount of raw data, in addition to other measurements (*e.g.* HR). Moreover, basic summary statistics of power such as session-averaged PO make practitioner to miss out acute changes of PO, yet being frequent in racing. Hence, it may give only a little insight into the real stress induced by the training or racing session. On this basis, some alternatives have been commonly provided. First, quantifying the total mechanical work would inform about the overall energy demand of the session, but it neglects any intensity specific effects such as those mentioned in the previous sub-sections. Otherwise, one could obviously looks at PO distribution per intensity zones, alike the framework behind eTRIMP and luTRIMP. This approach has two drawbacks yet:

1. Comparing means of PO presupposes to deal with homogeneous sample sizes in order to identify subtle differences, whereas the time past at high intensity remains often few.
2. It neglects time past at each intensity zones, assuming that performing 5 minutes at 300 W induces the same stress as 30 minutes at the same PO.

Back to the TSS, a power-dependent intensity weighting factor IF tackles well the aforementioned issues, by accounting for the fact that the physiological stress induced by exercise depends on PO itself. The training stress score is defined in the sequel:

$$TSS = \frac{D\, P_{norm}\, IF}{3600\, FTP} 100\,.$$

Here, D denotes the exercise duration in seconds, P_{norm} is the normalised PO using a 30 seconds time bins rolling average (a time close to the half-life of many physiological responses during and after exercise) (Coggan 2003), FTP is the functional threshold power such as a PO value that can be theoretically sustained over 60 minutes (prior determined through a specific test), IF is given by dividing P_{norm} by FTP. Since the TSS is normalised to the individual's threshold power, it allows for comparing TSS across athletes. Furthermore, a classification of IF scores into five incremental levels may arise. The first level has $IF < 0.75$ and describes a low intensity – recovery session, the second with $0.75 \leq IF < 0.85$ describes endurance training sessions, the third for $0.85 \leq IF < 0.95$ relates intermittent exercises at moderate intensities. The fourth level is described either by intervals performed at high intensities and short races with $0.95 \leq IF < 1.05$, or by shorter time trials ($1.05 \leq IF < 1.15$). Finally, the last level represents all-out exercises and track pursuit with $IF > 1.15$ (Coggan 2003).

Training stress score being a generalisable parameter, many adaptations have been provided (Coggan 2003). It makes such TL quantification method not only valuable for cycling, but also for running (Van Dijk and Van Megen 2017) and swimming activities.

1.1.1.2. Quantification in resistance training

While endurance sports –and more generally activities that are based on human locomotion– benefit from HR and any wearable sensor that captures the slightest exercise, monitoring the exercise demand objectively in resistance training (RT) remains challenging so far. As aforementioned, physiological status may inform how the athlete responds to exercise. This implies to regularly measure some physiological parameters of importance for a given population (amateurs or elite athletes), with all the underpinning methodological issues (Davison, Van Someren, and Jones 2009) to such assessment.

Among the physiological parameters of interest, hormonal responses to RT have been widely studied so far (Häkkinen and Pakarinen 1995; Hoffman, Im, Rundell, et al. 2003; Kraemer 1987; Kraemer, Fleck, Dziados, et al. 1993; Kraemer, Häkkinen, Newton, et al. 1998; Kraemer, Marchitelli, Gordon, et al. 1990; Kraemer and Ratamess 2005; Pullinen, Mero, Huttunen, et al. 2002; Shaner, Vingren, Hatfield, et al. 2014; Smilios, Pilianidis, Karamouzis, et al. 2003; Spiering, Kraemer, Anderson, et al. 2008; Vingren, Kraemer, Ratamess, et al. 2010). A narrative review conducted by Kraemer and Ratamess 2005 highlighted the most common hormonal acute responses and chronic adaptations to RT.

At first glance, testosterone may be of primary interest in a monitoring process according to its potential effects on force production enhancement. Resistance exercises can significantly increase testosterone concentrations in men through several factors (Kraemer and Ratamess 2005). Exercise selection in terms of the amount of muscle mass solicited and the poly-articular nature of exercise significantly impact acute testosterone elevations (Hansen, Kvorning, Kjaer, et al. 2001). Interaction between volume and intensity parameters drives acute changes in testosterone, with a predominance on the exercise volume. Indeed, training sessions that yield to a higher glycolytic component (e.g. characterised by a moderate intensity, high volume and short inter-set recovery duration) seem to provide the most considerable acute change (Crewther, Cronin, Keogh, et al. 2008; Kraemer, Gordon, Fleck, et al. 1991; Kraemer, Marchitelli, Gordon, et al. 1990). Furthermore, elevation in insulin concentration such as observed through high carbohydrate dietary intake showed a decreased circulating concentration of testosterone (Chandler, Byrne, Patterson, et al. 1994; Volek, Kraemer, Bush, et al. 1997). It implies that nutritional intakes may significantly impact testosterone changes and support the fact that the nutritional state is a leading parameter in any training process. Finally, chronic changes in testosterone following RT remain

inconsistent in the literature since resting levels are impacted by many factors, external to training (Kraemer and Ratamess 2005).

As for testosterone, concentrations in growth hormones (GH) appear to be elevated following RT in any population and share the same training related explanatory factors (e.g. volume, intensity, muscle mass recruited, rest intervals and training status). Among these, the total work performed seems to be one of the most determining factor of GH elevation (Hoffman, Im, Rundell, et al. 2003; Smilios, Pilianidis, Karamouzis, et al. 2003; Zafeiridis, Smilios, Considine, et al. 2003), where GH responses are correlated to $[lact_b]$ at and following exercise (Hakkinen and Pakarinen 1993; Kraemer and Ratamess 2005). However, not any change in resting values of GH concentrations are observed following traditional RT protocols. That supports the chronic modulator roles in the homeostatic functions of the physiological systems such as the regulation of glucose concentrations but also the acute roles in muscular remodelling process following RT (Kraemer and Ratamess 2005).

Levels of cortisol may be another parameter of interest for training monitoring. Briefly, cortisol stimulates lipolysis in adipose cells and takes place in tissue remodelling while enhancing protein degradation and inhibiting protein synthesis in muscle cells. The subsequent release of lipids and amino acids into circulation (Biolo, Maggi, Williams, et al. 1995) thus reflects an increase in cortisol levels, which becomes a useful parameter to be monitored through resistance training programs.

Following RT, acute responses of cortisol showed significant elevations along with adrenocorticotropic hormone –a stimulator of cortisol release (Di Blasio, Izzicupo, Tacconi, et al. 2014; Häkkinen, Pakarinen, Alen, et al. 1988; Kraemer, Fleck, Dziados, et al. 1993; Kraemer, Häkkinen, Newton, et al. 1998; Kraemer and Ratamess 2005; Kraemer, Fry, Warren, et al. 1992; McGuigan, Egan, and Foster 2004). Like GH responses, changes in cortisol are mostly observed following high volume RT performed at moderate intensities (Bottaro, Martins, Gentil, et al. 2009; Crewther, Cronin, Keogh, et al. 2008; Leite, Prestes, Rosa, et al. 2011; McCaulley, McBride, Cormie, et al. 2009; Smilios, Pilianidis, Karamouzis, et al. 2003). Accordingly, acute responses of cortisol showed positive associations with $[lact_b]$ elevation for such RT, usually performed for muscle hypertrophy purposes (Kraemer, Noble, Clark, et al. 1987; Ratamess, Kraemer, Volek, et al. 2005). Finally, inter-set recovery duration also weights the magnitude of post-exercise changes in cortisol concentrations.

Described by changes in resting cortisol concentration, chronic cortisol adaptations to RT are mitigated in the literature. Most of the studies reported by Kraemer and Ratamess 2005 showed either no changes or reductions in cortisol concentrations. Yet, some elevation in resting cortisol levels may occur after several weeks of RT (Häkkinen and Pakarinen 1991; Willoughby 2004). In the latter (Willoughby 2004), untrained people showed elevated cortisol concentrations after twelve weeks of RT along with significant increases of strength, body mass, fat-free

mass and muscle size gains. Authors conclude that the serum myostatin expression in response to cortisol elevation is not related to the training-induced muscle adaptations. This supports a more complex –at least non-linear– relationship that includes attenuating effects of follistatin-like related gene levels such as shown by a down-regulation of the activin IIb receptor.

Hormonal concentrations can be measured either in serum, plasma or saliva. In case of blood sample collection, the measures may be of great accuracy but also invasive and inappropriate for daily use (Helms, Kwan, Sousa, et al. 2020). In addition, sample collection and processing time raise the question of what pertinence can be ensured for such biochemical measures in a training monitoring process? Therefore, sports scientists and practitioners have the merit of striving to find out an indirect estimate of physiological stress induced by RT. In the sequel, we will present some of the main TL quantification methods used in RT.

Since RT can induce significant striated skeletal muscle adaptations, indirect biomarkers of muscle damage may be valuable in a monitoring process (Heckel, Atlasz, Tékus, et al. 2019; Helms, Kwan, Sousa, et al. 2020). Among the ones commonly measured, creatine kinase (CK), C-reactive protein and cytokines (e.g. Interleukines IL-1β and interleukines-6 (IL-6), tumour necrosis factorα) are potentially the best informative parameters of training effects on skeletal muscle structure.

Creatine kinase levels and IL-6 are known to raise a few days after training (Chen and HSIEH 2001; Nosaka and Newton 2002), in particular when exercise intensity is moderate to high (Baird, Graham, Baker, et al. 2012; Forti, Van Roie, Njemini, et al. 2017; Tiidus and Ianuzzo 1983) and when volume of exercises increases (Rodrigues, Dantas, Salles, et al. 2010; Tiidus and Ianuzzo 1983). However, it is important to consider the surrounding factors that influence biomarkers levels while attempting to quantify, partly, muscle damages induced by RE. Generally, changes in CK levels are a normal process, an indicator of energetic enzyme activity and muscle disturbances (Baird, Graham, Baker, et al. 2012). Beyond ethnic properties and gender on basal CK levels, hydration status before exercise may have a great influence on CK levels in response to exercise and lead to within-subject differences for comparisons (Fielding, Violan, Svetkey, et al. 2000). As an example among many, it may explain divergent findings in the literature about acute and chronic changes in such biomarkers due to uncontrolled confounding factors or training load parameters (Helms, Kwan, Sousa, et al. 2020). Finally, whatever the proxy marker of muscle damage to be measured, interpretation of these markers should not be done in isolation but rather along with other physiological, psychological, nutritional and environmental parameters (Helms, Kwan, Sousa, et al. 2020).

The first and basic training load quantification in RT comes with the volume load (VL) (Haff 2010; Scott, Duthie, Thornton, et al. 2016), a basic measure of amount of absolute loads lifted through a training session. Expressed in kilograms

or tonnes, it is simply calculated by the product of the number of sets, the number of repetitions and the absolute load such as:

$$VL_{abs} = V\ I_a. \tag{1.5}$$

In the equation, V denotes the product of the number of sets and the number of repetitions performed per set and I_a is the absolute load (kg).

The absolute VL is daily easy handling, but it presents several limitations. Firstly, using absolute values of intensity restricts the comparisons to within individuals comparisons only since athletes or practitioners have different strength levels. However, a simple adaptation of Equation 1.5 using relative measures of intensity rather than absolute would solve this issue. The relative VL is thus defined such as

$$VL_{rel} = V\ I_r, \tag{1.6}$$

with I_r being the load relative to the one repetition maximum (RM, %).

A major drawback to both absolute and relative VL is the fact that volume and intensity are reciprocal. For example, VL considers that performing N sets of M repetitions at 70% of 1RM equals performing M sets of N repetitions at the same load, which is obviously not realistic. In terms of training responses, such an assumption is theoretically incorrect due to the specific effects of exercise intensity on tissues (e.g. fibre types I and II hypertrophic responses Fry 2004), hormonal (e.g. growth hormone Ahtiainen, Pakarinen, Kraemer, et al. 2003; Vanhelder, Radomski, and Goode 1984 and cortisol responses Ahtiainen, Pakarinen, Kraemer, et al. 2003, chronic insuline-like growth factor-1, β-Endorphins and fluid regulatory hormones changes Kraemer and Ratamess 2005) and metabolic changes (e.g. blood lactate concentrations) Lagally, Robertson, Gallagher, et al. 2002.

Some authors attempted to solve this issue by planning sets of RE by holding some repetitions in reserve (i.e. shortening sets by a few repetitions before failure) (Genner and Weston 2014). This practice allows for practitioners to increase training volume compared to sets performed to failure. Aside from bringing subjective information to the calculation (subjective methods will be presented further), it may also result in a disproportionate reduction of volumes between sets. For example, performing 8 repetitions at a 10RM load results in a 20% volume reduction (according to the definition of volume given in Equation 1.5), whereas the same shortening over a 4RM set results in a 50% reduction of volume.

Another adaptation of the aforementioned VL, is based on RM prediction equations for a given number of repetitions (LeSuer, McCormick, Mayhew, et al. 1997) and 1RM predictions (Brzycki 1993). Defined in the sequel (see Equation 1.7), the athlete has a specific load for each RM of interest and consequently, an overall VL estimate that is not reciprocal and that varies across sets.

$$VL_{rm} = V\ I_{rr}. \tag{1.7}$$

Here, I_{rr} denotes the load relative to the RM for a given repetition range. However, the method relies on equations only appropriate for certain exercises, and which do not assume individual differences in the relationship between 1RM and repetitions that can be performed at a given load.

Despite these extensions of the former VL calculation sought to overcome some practical issues, it remains several limitations that may discredit the use of VL as an objective TL quantification method of *external* TL. Indeed, the travel of the load should be considered in any objective TL quantification, discarding the simplest formulations of VL where no differentiation can be made between resistance exercises. Otherwise, a rough depiction of the overall TL encountered may be expected when different resistance exercises are performed. One solution would include the barbell travelling or the centre of mass displacement in the equation, be it within an external mechanical work calculation (Fenn 1930; Marston, Peiffer, Newton, et al. 2017; McBride, McCaulley, Cormie, et al. 2009) or an extension of the original VL Haff 2010. For this purpose, RT benefits from the latest technological improvement that makes kinematic measures relatively simple through linear position transducer (LPT) and IMU (Pérez-Castilla, Piepoli, Delgado-Garcı+a, et al. 2019; Weakley, Morrison, Garcı+a-Ramos, et al. 2021).

Furthermore, the sessional intensity in which all exercises contribute is greatly affected by the session design. It would be safe to state that inter-set recovery duration impacts training outcomes in several ways Bird, Tarpenning, and Marino 2005. A systematic review conducted by Grgic, Schoenfeld, Skrepnik, et al. 2018 that included twenty-three studies for a total of 491 participants showed that inter-set recovery duration could impact the potential of muscular strength improvement in both amateurs and athletes. According to their findings, recovery durations over two minutes may allow for the greatest muscular strength gains. However, the influence of variability between recovery duration within a training session (which is common in practice) on training outcomes remains unclear. In terms of hormonal responses to volume equated RE, shorter is the recovery duration, greater is the magnitude of growth hormone (Boroujerdi and Rahimi 2008; Bottaro, Martins, Gentil, et al. 2009; De Salles, Simao, Miranda, et al. 2009; Rahimi, Qaderi, Faraji, et al. 2010) and serum cortisol acute responses (Rahimi, Ghaderi, Mirzaei, et al. 2010). Acute Blood lactate concentrations were also impacted by small rest intervals for a given load (Abdessemed, Duche, Hautier, et al. 1999; Kraemer and Ratamess 2005). As a final example, the energetic metabolism benefits from longer rest periods by better recovering of the ATP and PCr energy sources Harris, Edwards, Hultman, et al. 1976. Summing up, inter-set recovery duration has to be considered in the TL quantification in order to estimate better the response of the body to RT (Marston, Peiffer, Newton, et al. 2017).

Finally, none of the TL calculation methods presented so far considers neither the time that muscle is under tension or the repetition-velocity performed. Yet, time under tension (TUT) remains a key factor of the exercise response. It impacts muscle contractile properties, and yields to chronic neuromuscular adaptations

Burd, Andrews, West, et al. 2012; Tran, Docherty, and Behm 2006.

Beyond these specific issues related to methods for quantifying training loads, to prescribe training on RT percentage might also be problematic. To be the more possibly accurate, regular evaluations of 1RT should be performed since fluctuations in real RT operate through training cycles. In practice, such assessment may be time consuming and might interfere with the training prescription. In addition, the number of repetitions that can be performed at a given percentage of 1RM differs between individuals according to their own load-velocity (L-V) profiles (Banyard, Nosaka, Vernon, et al. 2018; Banyard, Nosaka, Sato, et al. 2017; Banyard, Tufano, Delgado, et al. 2018; Conceição, Fernandes, Lewis, et al. 2016; Garcı+a-Ramos, Pestaña-Melero, Pérez-Castilla, et al. 2018). From this basis, velocity-based training (VBT) comes as a contemporary method for training programming in respect of individual L-V profiles for a more accurate prescription of RT.

Velocity-based training comes as feedback to drive the practitioner on his movement or as an essential parameter integrated to the whole training prescription, alternatively to the traditional method based on 1RM percentage (Weakley, Mann, Banyard, et al. 2021). In the latter, determination of L-V profiles upon each movement of interest allows for prescribing absolute loads according to the individual L-V properties. In doing so, the daily variability in 1RM would not be an issue anymore since the velocity at each 1RM percentage seems to be stable enough (Banyard, Nosaka, Vernon, et al. 2018). The velocity of execution itself could thus be used as a valuable indicator of the effort done by practitioners. It could be directly integrated into a RT monitoring process (Balsalobre-Fernández and Torres-Ronda 2021).

Velocity-based training implies collecting data from every single repetition over each exercise, using measurement systems such as LPT or IMU. These systems have the benefit of measuring forces produced at exercise, external mechanical power and work. These variables represent objective measures of intensity and volume of RE, and may be directly considered in any monitoring process. Finally, VBT has a particular financial cost, but accurate and ecological measures of the least effort may be worth it.

1.1.1.3. Quantification in field sports: contribution of micro-technology

Contemporary approach for measuring the exercise demand in field sports is also marked by the raise of wearable sensors. As they are mainly based on running and cycling exercises, we retrieve the physiological principles underpinning to exercise, described in Section 1.1.1.1. Naturally intermittent, HR-based TL estimates remain a valuable parameter to be measured all through a training session or a game. Nevertheless, the quantification of TL in field sports remains mostly based on *external* TL estimates.

Global navigation satellite system coupled with inertial measurement units such as

common sport GPS has become the gold standard in team sports practiced outdoor. For indoor team sports, local positioning systems (LPS) overtake infrastructure related issues using base stations and radio-frequency signal for communicating between athletes and the positions of reference. The literature has also extended the last decade, assessing concurrent validity and reliability (Coutts and Duffield 2010; Crang, Duthie, Cole, et al. 2020; Fischer-Sonderegger, Taube, Rumo, et al. 2021; Jennings, Cormack, Coutts, et al. 2010; Rico-González, Los Arcos, Clemente, et al. 2020; Scott, Scott, and Kelly 2016) of such wearable sensors and providing areas of application in terms of training quantification (Gómez-Carmona, Bastida-Castillo, Ibáñez, et al. 2020; Jones, West, Crewther, et al. 2015; Rago, Brito, Figueiredo, et al. 2020; Ravé, Granacher, Boullosa, et al. 2020; Varley, Jaspers, Helsen, et al. 2017) for monitoring purposes. However, a thorough knowledge of how GPS, LPS and IMU estimate position data are still necessary in order to consider data at their fair value (Varley, Jaspers, Helsen, et al. 2017). For example, speed-derived data are usually measured through GPS signal source, whereas the number of collision or direction changes are usually estimated through IMU (Howe, Aughey, Hopkins, et al. 2017). While IMU benefit from much more sampling frequency (over 100 Hz), more accurate estimate would be expected. Yet, the reason for the use of GPS data instead of only IMU though being potentially more precise for tracking an athlete position and estimating the exercise demand (Vanwanseele, De Beéck, Schütte, et al. 2020) is that using IMU continuously remain a computational challenge in order to get low measurement errors (Zihajehzadeh, Loh, Lee, et al. 2015). On the other side, the use of GPS for quantifying exercise demand over short distances covered at high speed, including sharp turns suffer from error rates, according to the relatively low GPS sampling frequency (Crang, Duthie, Cole, et al. 2020; Scott, Scott, and Kelly 2016; Varley, Fairweather, and Aughey1 2012).

From a training monitoring point of view, GPS/ LPS sensors with embedded IMU allow for quantifying exercise intensity and volume objectively. A systematic review from Crang, Duthie, Cole, et al. 2020 retraces the main intensity and volume parameters that are commonly measured through these wearable devices and their validity regarding systems of reference. Placed between scapulae and fixed in an anatomically adjusted harness, they allow for measuring exercise intensity through pace, running velocity, running acceleration, mechanical power, and metabolic power – an EE representation– derived from acceleration (Osgnach, Poser, Bernardini, et al. 2010). Velocity and acceleration are usually divided into several zones, individualised according to MAS tests or arbitrary set (Rago, Brito, Figueiredo, et al. 2020).

On another note, the volume is essentially characterised as a time spent at running or the total distance covered. The sessional TL may be estimated by the product of intensity and volume parameters, or through accelerometers derived features. In the latter, common features raised from IMU are Dynamic Stress Load (*i.e.* the total of weighted impacts, Statsports, Ireland) (Gaudino, Iaia, Strudwick, et al. 2015),

New Body Load (i.e. the accumulation of forces; Catapult, Australia) (Ehrmann, Duncan, Sindhusake, et al. 2016), Force Load (i.e. the sums the forces produced from both foot strikes and collisions) (Colby, Dawson, Heasman, et al. 2014) and PlayerLoad^{TM} (the sum of the norm of accelerations; Catapult, Australia) (Barrett, Midgley, and Lovell 2014; Boyd, Ball, and Aughey 2013; Nicolella, Torres-Ronda, Saylor, et al. 2018). To date, PlayerLoad remains one of the most used feature for quantifying a session TL (Gómez-Carmona, Bastida-Castillo, Ibáñez, et al. 2020). It is defined in the sequel:

$$PlayerLoad^{TM} = \sum_{t=0}^{T} \sqrt{\frac{\left(a_{x(t)} - a_{x(t-1)}\right)^2 + \left(a_{y(t)} - a_{y(t-1)}\right)^2 + \left(a_{z(t)} - a_{z(t-1)}\right)^2}{100}},$$

where a_x, a_y and a_z denote the acceleration in the antero-posterior, lateral and vertical axis. In any cases, these features are aggregated according to a summation of accelerometer signal and aim at representing a "biomechanical" constraint sudden by an athlete at exercise.

There is indeed interest for estimating concurrently physiological and biomechanical adaptations to exercise through GPS– IMU sensors, in particular for injury prevention purposes (Vanrenterghem, Nedergaard, Robinson, et al. 2017). However, the relationship between both adaptations is not straightforward since physiological and bio-mechanical changes often rely on different time-frames. That is important when we attempt to understand the body's response to exercise for future training prescriptions. As an example, let consider a mechanical adaptation such as musculo-articular tissue adaptation (*e.g.* muscle thickness, fascicle length, pennation angle, tendon stiffness) (Wisdom, Delp, and Kuhl 2015) and underpinning physiological changes to given mechanical stimuli (*e.g.* changes in cell volume, muscle glycogen availability). According to the super-compensation principle (Bompa and Haff 2009; Viru 1984), injury may occur when biomechanical properties raise a critical point. At the same time, physiological states may have returned to a baseline level due to a shorter time course. In brief, different periodisation should be planned according to on one side, bio-mechanical and on the other side physiological stresses induced by exercise (Vanrenterghem, Nedergaard, Robinson, et al. 2017).

Finally, wearable GPS sensors allow for linking a planned TL to what TL has really been achieved by athletes and how the upcoming training should be adjusted regarding the collected data. However, rigour guidelines exist in order to ensure robustness in data collection and meaningful interpretation drawn for data. For example, each athlete should wear its own sensor to limit variability in the measurement and mistakes in data collection. In addition, these sensors do not measure everything. Tactics, environment, psychological and nutritional athlete status should be concurrently considered when interpreting training and game

outcomes from data.

Objective measures of TL are valuable for accurately estimating stresses imposed on the body. Measuring *internal* (i.e. physiological) stresses is not always possible and give away to *external* (i.e. physical) load measurements. Yet, relationships between physiological responses at exercise and training load in physical terms are not straightforward and may lead to misunderstanding how an athlete responds to a particular training session. For that reason, subjective measures of TL may bring additional and valuable information being appropriate in any context.

1.1.2. Subjective measures

1.1.2.1. Rate of perceived exertion

To date, acute body responses to training are alternatively and often estimated through the use of subjective rate of perceived exertion (RPE), as a substitute to the objective exercise intensity assessments. In the field of sports sciences, it has been initiated by Borg 1970. The first RPE came with the need for a single term that may represent how strenuous was the exercise or the sequence of exercises and how physiological systems respond to exercise. With that self-evaluation, physiological and psychological stresses are concurrently estimated to evaluate an overall response to exercise and allow further inter-individual and inter-process comparisons. Such an indirect measure of exercise difficulty (which combines intensity as well as exercise duration) was also intended to be part of any training monitoring process, in replacement or additionally to other objective measures of exercise intensity. Ratings of perceived exertion had to be theoretically supported by physiological responses to exercise. However, they also had to be robust, applicable and generalisable to any training situations (*i.e.* to any intensity levels, exercise location, levels of expertise and more generally to any sports without necessitating any specific material).

Some scales have been provided since the first formulation of RPE (Borg 1970). The former –so-called Borg RPE scale– was designed to grow linearly with HR and exercise intensity during incremental exercises ($r \in [0.8, 0.9]$)(Borg 1970, 1985, 1998). It came with a numerical fifteen graded-scale that ranges from 6 to 20 (see Table 1.1 for details) and allows for intra-individual numerical interpretation and comparisons of how difficult and strenuous was the exercise. Translated in many languages, that scale remains very popular so far.

Table 1.1. – Original Borg rating of perceived exertion scale by Borg 1998

Rating	Descriptor
6	No exertion at all
7	Extremely light
8	
9	Very light
10	
11	Light
12	
13	Somewhat hard
14	
15	Hard (heavy)
16	
17	Very hard
18	
19	Extremely hard
20	Maximal exertion

A few years later, a 0-10 category-ratio scale (CR10) has been provided by Borg 1990, described in Table 1.2. Scaling allows using more generalisable scales through several domains (e.g. physics, medicine), including an absolute zero and equivalent distances between levels. In the CR10, intensity descriptors are appropriately anchored to a particular number making the scale valuable not only for inter-individual but also inter-modal comparisons (e.g. stress, pain, difficulty) in any application and population.

Unlike the Borg RPE scale, the CR10 provides a non-linear growth function of intensity characterised by a positively accelerating increase such as an exponential curve (Borg 1990; Noble, Borg, Jacobs, et al. 1983). On the psycho–physiological side, CR10 was better correlated with accumulation of muscle lactate concentrations at exercise than the Borg RPE scale, the latter being more appropriate in regards to HR responses and RPE relationship. The CR10 has thus become a reference in RPE scales since it appears to be more appropriate for maximal, supra-maximal and intermittent exercises Borg and Borg 2010.

Table 1.2. – Original category-ratio scale of perceived exertion by Borg 1990

Rating	Descriptor	Attribute
0	Nothing at all	
0.5	Very, very weak	just noticeable
1	Very weak	
2	Weak	light
3	Moderate	
4	Somewhat strong	
5	Strong	heavy
6		
7	Very strong	
8		
9		
10	Very, very strong	almost max
.	Maximal	

One drawback with the CR10 is that it may be a *too simple* scale for assessing the stress induced by exercise properly due to a limited number of possible scores. For that reason, an extension of CR10 scale, the "centiMax" or 0-100 "centimax" scale (CR100) scale, has been further developed (Borg 1998; Borg and Borg 2010) while conserving the non-linear function but improving the number of possible ratings. To date, the choice of the scale to be used partly remains of personal conviction since both CR10 and CR100 scales appear to be inter-changeable (Fanchini, Ferraresi, Modena, et al. 2016) and commutable (Borg and Borg 2010). However, no matter how accurate is the rating, using a larger scale allows for more sensitive rates and, consequently, more variability when exploring psycho-physiological relationships. It is not a drawback for such, but it has to be taken into consideration when reporting or looking for correlations between physiological changes and RPE, or when exploring the reliability of various scales together (Scott, Black, Quinn, et al. 2013).

A sessional TL can be estimated through a valuable subjective measure of exercise intensity regarding psycho-physiological responses to exercise. Foster, Florhaug, Franklin, et al. 2001 initially proposed a modification of the CR10 scale (see Table 1.3) that only uses integers and provides modified verbal anchors to better reflect American idiomatic English. Using that scale, the product of the overall difficulty perceived by the end of a training session and the session duration in minutes let estimate session rate of perceived exertion (sRPE). In the case of resistance training, the overall RPE is preferably multiplied by the number of repetitions (Borresen and Lambert 2009). Alternatively, one may report RPE for each group of exercises, to be further added and multiplied by the total duration of the session. Still compared with HR based methods, using sRPE showed strong correlation with

TL quantification methods based on HR data (e.g. TRIMP) (Foster, Florhaug, Franklin, et al. 2001; Scott, Black, Quinn, et al. 2013). It makes sRPE a promising method for quantifying TL in any context and without any material requirements (Singh, Foster, Tod, et al. 2007).

Table 1.3. – Modification of the CR10 scale by Foster, Florhaug, Franklin, et al. 2001

Rating	Descriptor
0	Rest
1	Very, very easy
2	Easy
3	Moderate
4	Somewhat hard
5	Hard
6	
7	Very hard
8	
9	
10	Maximal

However, none method of TL quantification is perfect. The main limitations to the use of sRPE for training monitoring purposes are that RPE is mainly influenced by intensity rather than by volume. That means performing several repetitions or practising for a long time at low intensity is perceived to be easier than performing few repetitions at high intensities (Borresen and Lambert 2009; Sweet, Foster, McGuigan, et al. 2004). In addition, one may note that RPE already account for the time past at exercise in its definition. Multiplying RPE with either a time or a number of repetitions might bring noise and question the theoretical basis behind sRPE (Agostinho, Philippe, Marcolino, et al. 2015). Many factors contribute to the individual perception of the effort, such as personal traits, environmental conditions, psychological states, ventilation rate, neurotransmitter levels, hormone and substrate concentrations (Borresen and Lambert 2008). It might explain the poor reliability found in RPE measurements across standardised sessions (Scott, Black, Quinn, et al. 2013). The decrease in sRPE reliability might also be exacerbated by using larger scales for its calculation (Arney, Glover, Fusco, et al. 2019; Scott, Black, Quinn, et al. 2013; Wallace, Slattery, Impellizzeri, et al. 2014).

Rating of perceived exertion remains a gold standard for quantifying TL across various training sessions so far. It is commonly considered as an *internal* training load indicator since it aims to represent psycho-physiological responses to exercise (Bourdon, Cardinale, Murray, et al. 2017; Foster, Rodriguez-Marroyo, and De Koning 2017; Haddad, Stylianides, Djaoui, et al. 2017; Helms, Kwan, Sousa, et al.

2020; Lupo, Ungureanu, Frati, et al. 2019). Despite some lack of reliability and, in some way, a restrictive approach, its ease of use in any condition and application makes sRPE a valuable method to be accounted for in any training monitoring process, along with other objective and subjective training load indicators.

1.1.2.2. Athlete self-report measures

Another way to estimate the body's response to training came recently with the use of wellness and health questionnaires. They share the same advantages as sRPE such as being suitable in any sports, ease of use, as well as limitations (i.e. factors that influence the perception of exertion). In the meantime, they account for multiple items, mainly regarding mood states (e.g. tension, depression, fatigue, motivation).

A recent systematic review from Jeffries, Wallace, Coutts, et al. 2020 highlighted the most common athlete self-report measures (ASRM) for monitoring training responses. In their review, authors categorised instruments of ASRM into multiple items (Profile of Mood States (POMS); Recovery-Stress Questionnaire Athletes (RESTQ)) and single item instruments (single items or self-developed multiple-item measures with or without a combined score, including Hooper wellness items). The initial POMS questionnaire relies on the evaluation of six different mood dimensions (tension-anxiety, depression-rejection, anger-hostility, vigor-activity, fatigue-inertia and confusion-bewilderment), defined by seven to fifteen adjectives per mood factors (McNair, Lorr, Droppleman, et al. 1971). The mood score for each factor is given by the sum of the ratings obtained from adjectives. The RESTQ (Kallus 1995; Kallus and Kellmann 2016) is declined into several versions, including one for sports in which two versions were provided using 76 and 52 items, respectively. While POMS and RESTQ require attention among athletes in order to complete questionnaires, it might be at the expense of boredom, fatigue and frustration due to time-consuming and redundant questionnaires (Jeffries, Wallace, Coutts, et al. 2020). On the opposite, single items ASRM comes with a more suitable, reproducible rating, preserving athlete buy-in in the monitoring process and ensuring good ratings. Therefore, it is unsurprisingly that single items appear to be the most practical ASRM system for athletes or teams who attend to collect measures daily.

The so-called wellness items stem from the work of Hooper et al. (Hooper and Mackinnon 1995; Hooper, Mackinnon, Howard, et al. 1995) for monitoring undertraining and overtraining, including the first four items (fatigue, sleep quality, muscle soreness and stress) and then four additional items (enjoyment of training, irritability, health causes of stress and unhappiness). The authors supported the first four items by highlighting correlations between disturbances in mood state and physiological changes in swimmers (Hooper, Mackinnon, Howard, et al. 1995). A Hooper index or score has also been further provided, such as the summation of each rating using a 1–7 levels Likert scale (Hooper, Mackinnon, Howard, et al.

1995). Over time, several adaptations or the original wellness grid were provided, mainly without any additional validation. Plenty of appellations were also used for referring to the wellness items, which might have brought some confusion about the wellness definition (Jeffries, Wallace, Coutts, et al. 2020).

To conclude, ASRM are valuable methods for a training load monitoring purpose. The main advantage of using ASRM and psychometric questionnaires is that they do not require any specific material and can be used in any situation. Basically of easy use, recent software development makes ASRM daily collectable using smartphones or tablet computers, without bringing not any time constraint to staff and athletes. On this basis, monitoring TL is conceivable through a combination of objective and subjective measures, or *external* and *internal* TL estimates. This opens up the possibility for exploring the effects of training on athlete performance and injury in a systemic way, by accounting for both psychological and physiological responses to exercise (Burgess 2017). As a key point, be the TL measures of objective or subjective nature, they should always be collected and appreciated together including various training load indicators.

1.2. Modelling the effects of training

Previously, we described the primary methods for quantifying training loads with their benefits and drawbacks. As a final note, aiming at understanding the relationship between training and athletic performance is a matter of a systemic issue since athletic performance is a complex system in which many systems interact with each other (Bazyler, Abbott, Bellon, et al. 2015). Hence, two goals arise from the modelling with:

1. To describe training effects on performance using interpretable models.
2. To make accurate predictions from these models, intended for optimising future training protocols for sports performance

The modelling comes from various scientific domains such as biology, physics, statistics, and computer science from this statement.

In this section, we sequentially present the main models used for elucidating the effects of training on athletic performance. Initiated by physiologists and further empowered by statisticians, modelling of athletic performance is becoming increasingly attractive for scientists, be they intend to advance the theoretical or applied research.

1.2.1. A collection of Fitness-Fatigue models

The first models of training effects were developed in the seventies, notably with the work of Banister, Calvert, Savage, et al. 1975 and Calvert, Banister,

Savage, et al. 1976. Initially, Calvert, Banister, Savage, et al. 1976 designed a system model of the effects of training on physical performance using a general transfer function that applies on various inputs. Being a multi-component model in which determinants of physical performance such as endurance, strength, skill and psychological factors were accounted for as input of the system, it would allow for modelling a physical performance in any sports. However, this system model requires to measure training according to each of the aforementioned factors. It more reflects a mere system model rather than an applicable one that would be valuable in practice. In addition, the model only relies on features that affect positively performance outcomes, neglecting all the negative and inevitable but transient psycho-physiological states that occur in response to training. As a consequence, it does not consider any limit of performance that could be reachable, making its progression somehow infinite which is obviously unrealistic.

A simplification of that system model of training effects on performance is the so-called FFM (Calvert, Banister, Savage, et al. 1976). The model came with a single input but two transfer functions providing two antagonistic features (fitness and fatigue states) that affect positively and negatively performance outcomes, respectively. Any training session induces a positive long-lasting and low magnitude fitness effect and a negative short-lasting and high magnitude fatigue effect. The Fitness-Fatigue model is tailored to be fitted per individual, according to the singularity of training responses among athletes. Usually, fitness-fatigue models (FFMs) are fitted on historical data (*i.e.* retrospectively) for describing and predicting the effects of training on physical performance. Fitness-fatigue models would allow for optimising training protocols in order to enhance the greatest physical performance, using analytical solutions given by the system (*i.e.* the optimal "training dose" and the optimal rest time between consecutive training bouts). It is a matter of seeking the optimal command for adjusting states variables to reach the highest athletic performance. Some authors resumed how FFM could be used for that purpose by predicting performances from a given training load and conversely and by finding the training load for a given athletic performance (see Figure 1.2) (Busso and Thomas 2006). Therefore, modelling the effect of training using FFMs would provide some insights about how programming training cycles should be constructed within a given timeframe, including quantitative progressions of sessional training loads through overloading and tapering cycles.

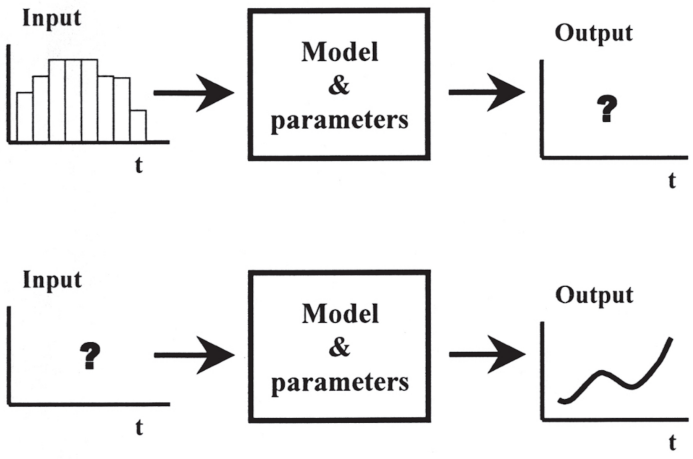


Figure 1.2. – Schematic representation about the use of FFM for making predictions and determining the optimal command , taken from Busso and Thomas 2006.

Several limitations of FFMs regarding the model stability, parameter interpretability, ill-conditioning and predictive accuracy have been reported (Hellard, Avalos, Lacoste, et al. 2006; Ludwig, Asteroth, Rasche, et al. 2019). For that reason, some extensions of FFMs have been further provided, attempting to improve the performances of the original FFM framework (Busso 2003; Busso, Denis, Bonnefoy, et al. 1997; Kolossa, Azhar, Rasche, et al. 2017; Ludwig, Asteroth, Rasche, et al. 2019). System models of training effects on performance thus relate to a collection of FFMs rather than a unique model.

1.2.1.1. Original fitness-fatigue model: a two-components model

In its original form, FFM framework relies on first-order differential equations (see Figure 1.3) and constitutes a linear time invariant model (see Equation 1.8). It allows for estimating a performance at any time t, according to fitness and fatigue responses to a training bout.

$$\hat{y}_{ban}(t) = \alpha_0 + k_g g(t) - k_h h(t) \quad \text{with } g(t) \geq 0 \text{ and } h(t) \geq 0\,. \tag{1.8}$$

In Equation 1.8, α_0 denotes a basic level of performance (i.e. a model intercept). In practice, α_0 may be included as a model parameter to be optimised along with the scaling factors k_g and k_h and time constants, or arbitrary fixed based on expert knowledge. Fitness and fatigue states are given by the impulse responses $g(t)$ and $h(t)$, defined in the sequel (see Equations 1.13 and 1.14.

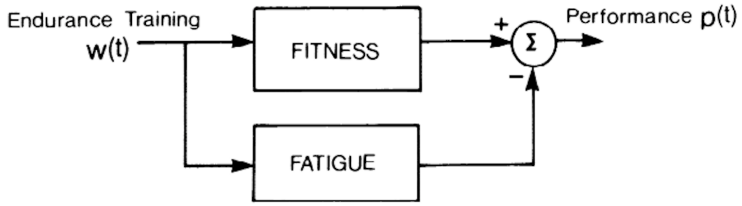


Figure 1.3. – Diagram representation of the two component Fitness-Fatigue model, taken from Morton, Fitz-Clarke, and Banister 1990

The linear system relies on convolution based features, in which each training input defined by a discrete function $w(t)$ is convolved with two impulse responses $G(t)$ and $H(t)$ (see Equations 1.9 and 1.10). Historically, $w(t)$ takes the form of bTRIMP (see Equation 1.2) since the model was developed along with endurance athletes, but it may take any form as long as it represents a sessional TL.

$$G(t) = e^{\frac{-t}{\tau_g}} \quad and \tag{1.9}$$

$$H(t) = e^{\frac{-t}{\tau_h}}. \tag{1.10}$$

The dynamical system of fitness and fatigue states responses to training bouts is defined in the sequel :

$$g'(t) = w(t) - \frac{1}{\tau_g} g(t) \quad and \tag{1.11}$$

$$h'(t) = w(t) - \frac{1}{\tau_h} g(t)\,, \tag{1.12}$$

where τ_g and τ_h denote the time constants for the fitness and fatigues states, such as $w(t) \geq 0$, $\tau_g \geq 0$ and $\tau_h \geq 0$. Consequently, a numerical solution of fitness and fatigue states is given from the convolution between the discrete function $w(t)$ and the impulse responses, such as :

$$\begin{aligned} g(t) &= w(t) * e^{\frac{-t}{\tau_g}} \\ &= \int_0^t w(s) e^{\frac{-1(t-s)}{\tau_g}} ds \quad and \end{aligned} \tag{1.13}$$

$$\begin{aligned} h(t) &= w(t) * e^{\frac{-t}{\tau_h}} \\ &= \int_0^t w(s)\, e^{\frac{-1(t-s)}{\tau_h}} ds\,. \end{aligned} \tag{1.14}$$

Here, $*$ denotes the convolution product between the discrete function $w(t)$ and the given impulse responses $g(t)$, $h(t)$. A complete mathematical derivation leading to Equations 1.11 and 1.13 in given in Appendix A. In the discrete form, Equations 1.13 and 1.14 may be rewritten, such as :

$$g(n) = \Delta t \sum_{i=1}^{n} w(i)\, e^{\frac{-(n-i)}{\tau_g}} \quad and \tag{1.15}$$

$$h(n) = \Delta t \sum_{i=1}^{n} w(i)\, e^{\frac{-(n-i)}{\tau_h}} \,. \tag{1.16}$$

Here, Δt denotes the discrete time step, usually one day. Hence, the model equation (see Equation 1.8) may be rewritten such as :

$$\hat{y}^{ban}(n) = \alpha_0 + k_g \sum_{i=1}^{n-1} w(i)\, e^{\frac{-(n-i)}{\tau_g}} - k_h \sum_{i=1}^{n-1} w(i)\, e^{\frac{-(n-i)}{\tau_h}} . \tag{1.17}$$

From the model equation, the time to recover performance t_r after a training session and to peak performance t_p after a the completion of training are explicitly given by Equations 1.18 and 1.19 (Fitz-Clarke, Morton, and Banister 1991).

$$t_r = \frac{\tau_g \tau_h}{\tau_g - \tau_h}\, ln\left(\frac{k_h}{k_g}\right) \quad and \tag{1.18}$$

$$t_p = \frac{\tau_g \tau_h}{\tau_g - \tau_h}\, ln\left(\frac{k_h \tau_g}{k_g \tau_h}\right) . \tag{1.19}$$

While the original system model seems to have proven its interest for training effect modelling purposes, it also presents some limitations. First, the greatest concerns with the framework presented above is both theoretical and practical. Given the formulation described in Equation 1.17, more an athlete trains, more he progresses in a infinite way (without reaching any limit).

Going deeper on the analysis or model performances, one might expect caveats about the predictive capability of FFM. Historically, FFM attempted to prove their performances based on the coefficient of determination R^2 (Hellard, Avalos, Lacoste, et al. 2006; Morton, Fitz-Clarke, and Banister 1990). That criterion informs about the amount of variance that is explained only by linear relationships between the target and the independent variables (i.e. the components of the FFM). Thus, it is not the most valuable criterion for assessing the model accuracy in predictive applications, since other criteria such as mean absolute error (MAE), MAPE or RMSE are usually recommended (Bartlett 1997; Chai and Draxler 2014; Willmott and Matsuura 2005).

A case study from Hellard, Avalos, Lacoste, et al. 2006 has evaluated the

goodness of fit, accuracy and ill-conditioning of the original FFM on historical data. Through bootstrap replicates, authors reported a reasonably high goodness of fit ($R^2 = 0.79 \pm 0.13$) and a MAE equal to $1.05 \pm 0.63\%$. However, collinearity between model parameters was observed, supporting an ill-conditioning problem. As a non-linear generalisation of the collinearity problem in linear regression, ill-conditioning may reveal an over-parametrisation of the model for a given data set (Bates and Watts 1988). This finding supports that complexifying the model (i.e. by adding first-order components to the model) does not necessarily improve the model performances in athletic performances modelling (Busso, Carasso, and Lacour 1991). In their study (Busso, Carasso, and Lacour 1991), authors reported lower and negative scaling factors k_h than ones of k_g for six of height subjects. It thus questions the rationale behind the use of the fatigue component in the model and more generally, the pertinence of complexifying the model in that way. However, one may note that these results apply to a particular data set and might not be replicable using another data set in other sports disciplines. Therefore, model validation processes should be performed in order to ensure that the FFMs employed in studies are of valuable structure, according to the data set used.

Fitness-Fatigue model suffers from its univariate configuration in modelling athletic performances. While it is known that athletic performance is multifactorial (Avalos, Hellard, and Chatard 2003; Bazyler, Abbott, Bellon, et al. 2015; Mujika, Busso, Lacoste, et al. 1996; Stone, Stone, and Sands 2007), variations in performances may not be fully explained by the dynamics of training loads only, resulting in poor predictive capability (Chiu and Barnes 2003; Pfeiffer 2008; Taha and Thomas 2003). Recently, Piatrikova, Willsmer, Altini, et al. 2021 provided a multivariate alternative to the original FFM. In their study, authors showed that using training loads and psychometric variables as a linear combination of model inputs yielded to better performing models. Yet, it implies that each input variable induces both fitness and fatigue responses, which might be questioned theoretically. Also, such results are not intended to be compared to ones from Busso, Carasso, and Lacour 1991 since the increase in model complexity is due to an addition of model inputs that stem from various sources, rather than an increase in model components for a given model input.

1.2.1.2. A linear time-varying Fitness-Fatigue model

To date, most of the studies used linear time-invariant FFM for modelling athletic performance (Banister, Calvert, Savage, et al. 1975; Banister and Hamilton 1985; Banister and Fitz-Clarke 1993; Banister, Green, McDougall, et al. 1991; Busso, Candau, and Lacour 1994; Busso, Carasso, and Lacour 1991; Busso, Häkkinen, Pakarinen, et al. 1990, 1992; Morton, Fitz-Clarke, and Banister 1990; Mujika, Busso, Lacoste, et al. 1996; Vermeire, Van de Casteele, Gosseries, et al. 2021). In these studies, authors considered that the model parameters (e.g. k_g, k_h, τ_g, τ_h) did not vary over time within athletes. To our knowledge, only two studies (Busso,

Benoit, Bonnefoy, et al. 2002; Busso, Denis, Bonnefoy, et al. 1997) provided a time-varying alternative by adjusting model parameters for each new entry, using a recursive least square algorithm (i.e. by minimising the residual sum of squares for each day of performance, such as defined in Equation 1.21). Basically, the residual sum of squares (RSS) is

$$RSS = \sum_{n=1}^{N} (\hat{y}_n - y_n)^2 \,, \tag{1.20}$$

where n denotes the day of an athletic performance y, $\hat{y}_n$ is the predicted performance. On day n, model parameters were fitted by minimising a recursive least square function $S(n)$, such as

$$S(n) = S(n-1)\,\alpha + (\hat{y}_n - y_n)^2 \,. \tag{1.21}$$

Here, α is a constant such as $0 < \alpha < 1$. Globally, re-calculating model parameters at each occurrence significantly improved the model performances by about 19% of explained variance in the model, when compared to a time-invariant model (Busso, Denis, Bonnefoy, et al. 1997). Also, one may notice the wide variability observed in the estimate of parameter coefficients (i.e. k_g, k_h, τ_g, τ_h), up to 108%. These results fill the FFM issues out exposed by Hellard, Avalos, Lacoste, et al. 2006 and ones from literature in which parameter estimates may greatly differ within and between athletes. In addition, authors (Busso, Benoit, Bonnefoy, et al. 2002; Busso, Denis, Bonnefoy, et al. 1997) attempted to draw interpretations from the variability in model parameters, in particular for the gain terms k_g and k_h. Such variations might either describe acute fatigue responses according to specific performances or artefacts from irregularities in the changes in performance over weeks (i.e. changes of performances that are not related to the training load dynamics) (Busso, Denis, Bonnefoy, et al. 1997).

In summary, a linear time-varying model appears to help model the effects of training on performance. In practical terms, it may highlight the influence of high-intensity training on both fitness and fatigue responses. That is important for attempting to model training effects on performance since the underpinning physiological mechanisms to exercise are complex and depend on surrounding training doses. However, fitting parameters recursively might provide a great variability in parameter estimates. Appropriate cautions should be taken when interpreting model parameters to optimise training prescriptions.

1.2.1.3. A preload concept for improving Fitness-fatigue models

Still based on the original FFM formulation (see Equations 1.8 and 1.17), issues were attributed the the intercept term α_0 –the basic level of performance–, notably with the work of Ludwig, Asteroth, Rasche, et al. 2019. In the sequel, we recall

three methods for estimating α_0, named $\alpha_{0,A}$, $\alpha_{0,B}$ and $\alpha_{0,C}$, respectively. Basically, $\alpha_{0,A}$ can be arbitrary set at a certain level of the first performance. For instance, authors (Busso 2003; Busso, Denis, Bonnefoy, et al. 1997) suggested to set the basic level of performance at 80% of the first known performance. It was justified by the fact that break periods may decrease the performance potential by about 20% (in this case, the study began right after a break period). Another way of setting value to $\alpha_{0,B}$ is to include α_0 as a n^{th} parameter to be fitted in the model (Chalencon, Pichot, Roche, et al. 2015; Clarke and Skiba 2013). Otherwise, $\alpha_{0,C}$ might be estimated from prior training history or arbitrary choosing a performance level if athletes are de-trained, according to Wood, Hayter, Rowbottom, et al. 2005. The last proposal seems to be the most valuable but it is often not possible, due to missing data prior the modelling period.
Setting α_0 such as $\alpha_0 = 0.8\, y_1$ or fitting α_0 as a model parameter have their own drawbacks and greatly influence the model performances. On one side, let consider a basic level of performance equal to 80% of the first known performance. In this case, if athletes are not adequately de-trained (e.g. by maintaining physical activity during break periods), the model will directly suffer from a biased α_0 estimate, leading to inaccurate predictions. Conversely, setting α_0 at a higher value than the *true* value (e.g. near the first known performance) results in a high lower bound. Accordingly, it implies a restricted range of possible values that could be taken by $\hat{y}$, where athletes could not reach a basic level of performance below α_0, even in the case of substantial de-training.

On this basis, a preload estimate provided by Ludwig, Asteroth, Rasche, et al. 2019 aims at solving that issue by estimating training effects that have been done prior the data collection. According to Equations 1.15 and 1.16, the general formulation of fitness and fatigue preload terms is

$$pr^{g}_{[-\infty:n]} = k_g\, e^{\frac{-x}{\tau_g}} \sum_{i=-\infty}^{n} w(i)\, e^{\frac{-(n-i)}{\tau_g}} \quad \textit{and} \tag{1.22}$$

$$pr^{h}_{[-\infty:n]} = k_h\, e^{\frac{-x}{\tau_h}} \sum_{i=-\infty}^{n} w(i)\, e^{\frac{-(n-i)}{\tau_h}} \,. \tag{1.23}$$

Here, the preload period is limited from $-\infty$ to the first known training n. The exponential scaling factor allows for the training effect to exponentially decay according to x, the number of day simulated minus the number of the first day included for simulation plus one such as $x \in \mathbb{N}$, and the time constants τ_g and τ_h for the fitness and fatigue states, respectively.
A more feasible version of the preload terms (see Equations 1.22 and 1.23) let estimate the preload over a defined period, so-called a "short-term" preload (Ludwig, Asteroth, Rasche, et al. 2019). It is defined in the sequel:

$$pr_n^g = k_g\, e^{\frac{-x}{\tau_g}} \sum_{i=1}^{n} w(i)\, e^{\frac{-(n-i)}{\tau_g}} \quad and \tag{1.24}$$

$$pr_n^h = k_h\, e^{\frac{-x}{\tau_h}} \sum_{i=1}^{n} w(i)\, e^{\frac{-(n-i)}{\tau_h}}\,. \tag{1.25}$$

Accordingly, the discrete formulation of the FFM including a preload term becomes

$$\hat{y}_{ban}(n) = (\alpha_0 + pr_{m-1}^g - pr_{m-1}^h) + k_g \sum_{i=m}^{n-1} w(i)\, e^{\frac{-(n-i)}{\tau_g}} - k_h \sum_{i=m}^{n-1} w(i)\, e^{\frac{-(n-i)}{\tau_h}}. \tag{1.26}$$

In practice, preload terms can be either computed considering unknown historical data or fitted on training data along with other FFM parameters, independently of the method for estimating the basic level of performance ($\alpha_{0,A}$, $\alpha_{0,B}$ and $\alpha_{0,C}$). Through data simulations and after having split data into training and test data set, Ludwig, Asteroth, Rasche, et al. 2019 showed that using fitted preloads on training data significantly improved the model accuracy when compared to no preloads, except for the use of $\alpha_{0,B}$ (α_0 fitted along with other parameters). However, predictions made using preload terms might result in a mean-like model, where the variance of predicted performance is close to null. Even if the fitting simulations benefit from better accuracy, no trend in modelled performances may be observed. When investigating the model performances using preload terms in test data, the authors did not report any significant improvement in predictions accuracy. However, it might be partly explained by the small sample size in the test samples.

In summary, using means of preload information that can be computed even if historical data are missing might considerably improve the model intercept estimate α_0. Since the intercept is a determining term in any linear model, further investigations that attempt to improve its estimate would be of particular interest.

1.2.1.4. A variable dose-response model

Resolving systems of linear differential equations such as done by original (Banister, Calvert, Savage, et al. 1975) and slight variants of FFMs that have been presented so far enabled to predict the training effects on physical performance. However, one of the main drawbacks of these models comes with the fact that they are based on the strong assumption that changes in performances depend exclusively on the current physical capacity, described by the fitness and fatigue states (or the fitness state only). It would thus be more appropriate to consider the influence of past training sessions on the states variables since they affect

physiological adaptations for several days (Dufaux, Assmann, Schachten, et al. 1982; Essig, Alderson, Ferguson, et al. 2000; Viru and Viru 2001). This hypothesis was first supported by the modification of the original FFM presented in Section 1.2.1.2, in which authors let model parameters to vary following each training bout (Busso, Benoit, Bonnefoy, et al. 2002; Busso, Denis, Bonnefoy, et al. 1997). Hence, authors made the assumption that state variable responses are not identical through training sessions, which probably better suits most of the physiological adaptations in human. In order to account for the past training sessions in the estimate of states variables, Busso 2003 proposed a gain term for the fatigue component that varies according to the previous training doses. According to Equations 1.9 and 1.10, the equation of the DR becomes

$$\hat{y}^{busso}(t) = \alpha_0 + k_g((w * G)(t) - ((k_h w) * H)(t)\,. \tag{1.27}$$

Here, $k_h(t)$ is related to the training doses by a second convolution to the transfer function

$$H_v(t) = e^{\frac{-t}{\tau_{h'}}}\,, \tag{1.28}$$

with $\tau_{h'}$ a time constant. Since it is defined as $k_h(t) = k_{h'}(w * H_v)(t)$ where $k_{h'}$ is a gain term, $k_h(t)$ increases proportionally to the training load and decay exponentially from this new value. In its discrete form, Equation 1.27 can be rewritten as

$$\begin{aligned} \hat{y}^{busso}(n) &= \alpha_0 + k_g \sum_{l=1}^{n-1} w(l)\, e^{\frac{-(n-l)}{\tau_g}} - \sum_{l=1}^{n-1} k_h(l)\, w(l) e^{\frac{-(n-l)}{\tau_h}}\,, \quad n \in \mathbb{N} \\ \text{with} \quad k_h(l) &= k_{h'} \sum_{m=1}^{l} w(m) e^{\frac{-(l-m)}{\tau_{h'}}}\,. \end{aligned} \tag{1.29}$$

As usual, the five parameters of the model (i.e. k_g, $k_{h'}$, τ_g, τ_h, $\tau_{h'}$) are fitted by minimizing the RSS (see Equation 1.20) between modelled and observed performances.

In addition and according to Equations 1.18 and 1.19, an optimal daily training load (ODT) allowing for maximal theoretical performance gains may be determined in the sequel:

$$ODT = \frac{k_g\, e^{\frac{-1}{\tau_g}} \left(1 - e^{\frac{-1}{\tau_h}}\right) \left(1 - e^{\frac{-1}{\tau_{h'}}}\right)}{2k_{h'} \left(1 - e^{\frac{-1}{\tau_g}}\right) e^{\frac{-1}{\tau_h}}} \tag{1.30}$$

In the study of Busso 2003, the proposed model provided a better fit –according to the adjusted R^2– than the FFMs using either one component (i.e. fitness component only) or two-components (fitness and fatigue components) (Busso 2003). Also, time-varying fatiguing effect (i.e. the fatigue state variable) seemed to improve

the DR model fit significantly, whereas it was not systematically true for the two-components and three-components FFMs using time-invariant parameters, in regard to the simpler formulation (one-component FFMs). This supports the fact that an optimal model complexity exists (*i.e.* a model that leads to the greatest model performances while being parsimonious), where complexifying FFMs would not systematically enhance better model performances and might be at the expense of model interpretations (Busso 2003; Busso, Carasso, and Lacour 1991).

1.2.1.5. Fitness-fatigue and delayed effects

While FFMs collections have been broaden by outstanding extensions (Busso 2003; Busso, Denis, Bonnefoy, et al. 1997), two recent modifications of the former FFM framework were proposed by (Matabuena and Rodrı+guez-López 2019). Authors shared the same objective as Busso 2003, attempting to make the most of past training sessions for estimating a physical performance. The first proposal relies on a FFM for delay differential equations. Formally, we consider a system of delay differential equations :

$$\begin{cases} g'(t) = w(t) - \frac{1}{\tau_1}g(t) - \frac{1}{\tau_2}g(t-1) - \cdot - \frac{1}{\tau_k}g(t-k) \\ h'(t) = w(t) - \frac{1}{\tau_{k+1}}h(t) - \frac{1}{\tau_{k+2}}h(t-1) - \cdot - \frac{1}{\tau_{2k}}h(t-k)\,. \end{cases} \tag{1.31}$$

Here, we suppose initial conditions $g = h = 0 \in [-k; 0]$, where k denotes the order of each delays and τ_i is a positive time constant with $i = 1, \cdot, 2k$. The discrete function $w(t)$ denotes the training load at time t.
According to Equation 1.31, we provide the solutions to the system of differential equations with one or three delays. The choice of extending the model to one and three delays was motivated by the following reasons:

1. Changes in physical condition are mainly influenced by 72 hours (Viru and Viru 2001).
2. Each delay term increases the number of parameters by two, so it might yield to highly complex model for usual small sample sizes and therefore impair model performances.

Model with one delay.

According to Equation 1.31, we rewrite the differential equation

$$g'(t) = -\frac{1}{\tau_1}g(t) - \frac{1}{\tau_2}g(t-1) + w(t)\,, \tag{1.32}$$

with initial conditions $g = 0$ on $[0; 1]$. Analogously, Equation 1.32 is equivalent to

$$g(t)e^{\frac{1}{\tau_1}(t-k)} - g(k) = \int_k^t \left[w(s) - \frac{1}{\tau_2}g(s-1)\right] e^{\frac{-1}{\tau_1}(s-k)} ds\,, \tag{1.33}$$

for $t \in [k; k+1],\ k = 0, 1, \ldots$

According to Equation 1.33, we have

$$g(t) = g(k)e^{-\frac{1}{\tau_1}(t-k)} + \int_k^t \left[w(s) - \frac{1}{\tau_2}g(s-1)\right]e^{-\frac{1}{\tau_1}(t-s)}ds\,,$$

for $t \in [k, k+1],\ k = 0, 1, \ldots$ and the following approximation:

$$\begin{aligned} g(k+1) &= g(k)e^{-\frac{1}{\tau_1}} + \left[w(k) - \frac{1}{\tau_2}g(k-1)\right]e^{-\frac{1}{\tau_1}} \\ &= \left[w(k) + g(k) - \frac{1}{\tau_2}g(k-1)\right]e^{-\frac{1}{\tau_1}}\,,\ k = 0, 1, \ldots \end{aligned}$$

For $t \in [0; 1]$, we have

$$\begin{aligned} g(t) &= \int_0^t \left[w(s) - \frac{1}{\tau_2}g(s-1)\right]e^{-\frac{1}{\tau_1}(t-s)}ds \\ &= \int_0^t w(s)e^{-\frac{1}{\tau_1}(t-s)}ds\,, \end{aligned} \tag{1.34}$$

which coincides with Equation 1.13.

In order to reach a discrete formulation of $g(t)$, Equation 1.32 may be rewritten such as

$$g(t) = \int_0^t \left[w(s) - \frac{1}{\tau_2}g(s-1)\right]e^{-\frac{1}{\tau_1}(t-s)}ds\,, \quad t \geq 0\,.$$

Therefore, we write

$$g(n) = \int_0^n \left[w(s) - \frac{1}{\tau_2}g(s-1)\right]e^{-\frac{-1}{\tau_1}(n-s)}ds\,, \quad t \geq 0\,,$$

such that can be approximated by the following equation

$$g(n) = \sum_{i=0}^{n-1} \left[w(i) - \frac{1}{\tau_2}g(i-1)\right]e^{\frac{-1}{\tau_1}(n-i)}\,.$$

If we consider that $g(-1) = 0)$ and so $w(0) = 0$, the equation becomes

$$g(n) = \sum_{i=1}^{n-1} \left[w(i) - \frac{1}{\tau_2}g(i-1)\right]e^{\frac{-1}{\tau_1}(n-i)}\,. \tag{1.35}$$

The full derivation from Equation 1.33 is displayed in Appendix B.

Model with three delays.

Using the formalism of the one delay model, authors provided an extended model with three delays in order to account for the effects of more past training sessions. In brief, according to Equation 1.32, we can consider a differential equation

$$g'(t) = -\frac{1}{\tau_1}g(t) - \frac{1}{\tau_2}g(t-1) - \frac{1}{\tau_3}g(t-2) - \frac{1}{\tau_4}g(t-3) + w(t)\,,$$

with initial condition $g = 0 \in [-3; 0]$. Its solution can be defined such as

$$\begin{aligned} g(t) &= \int_0^t \left[w(s) - \sum_{k=1}^{3} \frac{1}{\tau_{k+1}} g(s-k)\right] e^{-\frac{1}{\tau_1}(t-s)} ds\,, \quad t \geq 0 \\ &= g(k)e^{-\frac{1}{\tau_1}(t-k)} + \int_k^t \left[w(s) - \frac{1}{\tau_2}g(s-1) - \frac{1}{\tau_3}g(s-2) - \frac{1}{\tau_4}g(s-3)\right] e^{-\frac{1}{\tau_1}(t-s)}\,. \end{aligned} \tag{1.36}$$

According to Equations 1.35 and 1.36, its discrete form becomes

$$g(n) = \sum_{i=0}^{n-1} \left[w(i) - \frac{1}{\tau_2}g(i-1) - \frac{1}{\tau_3}g(i-2) - \frac{1}{\tau_4}g(i-3)\right] e^{-\frac{1}{\tau_1}(n-i)}\,. \tag{1.37}$$

Therefore, we can resolve the problem with the following approximation :

$$\begin{aligned} g(k+1) &= g(k)e^{-\frac{1}{\tau_1}} \left[w(k) - \frac{1}{\tau_2}g(k-1) - \frac{1}{\tau_3}g(k-2) - \frac{1}{\tau_4}g(k-3)\right] e^{-\frac{1}{\tau_1}} \\ &= \left[w(k) + g(k) - \frac{1}{\tau_2}g(k-1) - \frac{1}{\tau_3}g(k-2) - \frac{1}{\tau_4}g(k-3)\right] e^{-\frac{1}{\tau_1}}\,. \end{aligned} \tag{1.38}$$

for $k = 0, 1, \ldots$

Model based on an integral differential equation.

A more general formulation of the problem based on differential equation using an integral has also been introduced by Matabuena and Rodrı+guez-López 2019, intended for considering past information in order to estimate fitness and fatigue states. Briefly, it relies on a function that decreases in current time t and past times s. The integral differential equation is presented in the sequel :

$$g'(t) = -\frac{1}{\tau_1}g(t) - \frac{1}{\tau_5}\int_0^t \delta(t,s)g(s)ds + w(t)\,.$$

Here, τ_1 and τ_5 are positive time constants. We can consider that $\delta(t,s) = 0$ for $s < t-4$, so the function $g(t)$ becomes

$$\begin{aligned} g(t) &= g(0)e^{-\frac{1t}{\tau_1}} + \int_0^t \left[w(u) - \frac{1}{\tau_5}\int_0^u \delta(t,s)g(s)ds\right]e^{-\frac{t}{\tau_1}(t-u)}du \\ &= \int_0^t \left[w(u) - \frac{1}{\tau_5}\int_0^u \delta(t,s)g(s)ds\right]e^{-\frac{t}{\tau_1}(t-u)}du\,, \quad t \geq 0\,. \end{aligned} \tag{1.39}$$

Therefore, we can approximate $g(t)$ such as

$$\begin{aligned} g(n) &= \sum_{i=0}^{n-1}\left[w(i) - \frac{1}{\tau_5}\int_0^i \delta(i,s)g(s)ds\right]e^{-\frac{t}{\tau_1}(n-i)} \\ &= \sum_{i=0}^{n-1}\left[w(i) - \frac{1}{\tau_5}\sum_{j=0}^{n-1}\delta(i,j)g(j)\right]e^{-\frac{t}{\tau_1}(n-i)}\,. \end{aligned}$$

Then, we can assign fixed values to $\delta(i,i-1)$, $\delta(i,i-2)$ and $\delta(i,i-3)$. For instance, let define fixed values such as

$$\delta(i,i-1) = 0.5 \quad \delta(i,i-2) = 0.3\,, \quad \delta(i,i-3) = 0.2\,.$$

According to Equation 1.39 and for $t \in (k,k+1]$, $k = 0,1,\ldots$, we have

$$\begin{aligned} g(k+1) &= g(k)e^{-\frac{1}{\tau_1}} + \left[w(k) - \frac{1}{\tau_5}\int_0^k \delta(k,s)g(s)ds\right]e^{-\frac{t}{\tau_1}} \\ &= \left[w(k) + g(k) - \frac{1}{\tau_5}\sum_{j=k=3}^{k-1}\delta(k,j)g(j)\right]e^{-\frac{t}{\tau_1}} \\ &= \left[w(k) + g(k) + 0.5 - \frac{1}{\tau_5}g(k-1) + 0.3 - \frac{1}{\tau_5}g(k-2) + 0.2 - \frac{1}{\tau_5}g(k-3)\right]e^{-\frac{1}{\tau_1}}\,, \end{aligned} \tag{1.40}$$

for the given fixed values and for $k = 0,1,\ldots$. In this manner, we obtain similar results than the three delay approach using similar fixed values in Equation 1.38, such as $-\frac{1}{\tau_2} = 0.5 - \frac{1}{\tau_5}$, $-\frac{1}{\tau_3} = 0.3 - \frac{1}{\tau_5}$ and $-\frac{1}{\tau_4} = 0.2 - \frac{1}{\tau_5}$.

To conclude with these FFM extensions, authors showed promising results using data from (Clarke and Skiba 2013) and reached values of $R^2 = 0.99$ while fitting

their model on data from a single cyclist. These extensions remain to be further validated on other data set, but might provide a more realistic representation of bio-physiological kinetics involved by exercise (Miller, Olesen, Hansen, et al. 2005; Philippe, Borrani, Sanchez, et al. 2019; Viru and Viru 2001) than the original FFM from Banister, Calvert, Savage, et al. 1975.

1.2.1.6. Improving fitness-fatigue models with Kalman filter

Beyond the reformulation of the original FFM from Banister, Calvert, Savage, et al. 1975 presented so far, the use of KF (Kalman 1960; Welch, Bishop, et al. 1995) within FFMs has also showed great interest (Kolossa, Azhar, Rasche, et al. 2017). Having numerous applications in domains such as aeronautics, mechanical and electrical engineering, and neurosciences (Wolpert and Ghahramani 2000), the KF is an integral part of statistics and control theory. Briefly, KF is an optimal recursive data processing algorithm that processes a series of available measurements, regardless of their precision, and produces estimates of the variables of interest using (Maybeck 1990):

1. Knowledge of measurement systems.
2. Statistical description of system noises, uncertainty and measurement errors.
3. any available information about initial conditions.

Since FFMs are mostly univariate models that do not consider any other performance factors than training loads dynamics (e.g. psychological, nutritional, technical and environmental factors) (Hellard, Avalos, Lacoste, et al. 2006), a KF might improve the estimate of fitness and fatigue states variables using such optimal feedback of the prediction error. In this way, the un-modelled effects of training (e.g. psychological changes) or factors of athletic performance that are not accounted for are considered as state noise. In addition, one can assume measurement errors which might also bring noise in data, called observation noise. To date, only Kolossa, Azhar, Rasche, et al. 2017 have used a KF for optimising the predictive capability of FFMs. In their study, authors (Kolossa, Azhar, Rasche, et al. 2017) used the Busso's linear time-varying extension of the original FFM (Busso 2003) presented in Section 1.2.1.4. The implementation of a KF within the DR model is shown in Figure 1.4. A mathematical description of FFM combined with a KF is provided in Appendix C.

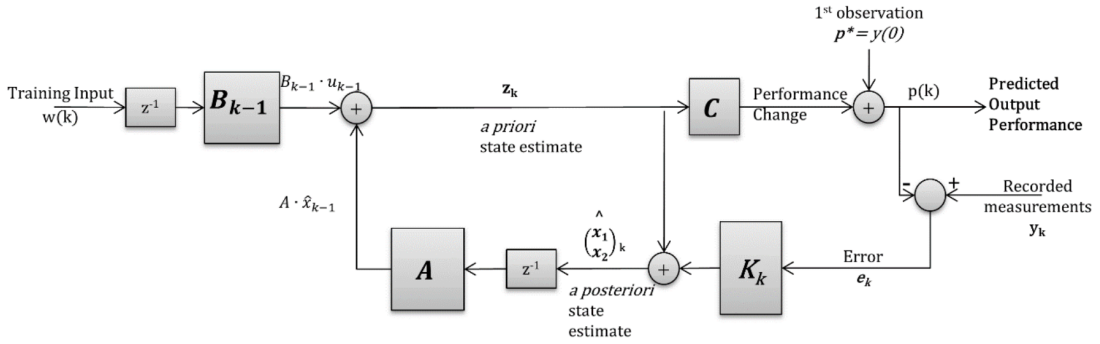


Figure 1.4. – Block diagram of a KF system, taken from Kolossa, Azhar, Rasche, et al. 2017. $\mathbf{B_{k-1}}$ is the time-varying input matrix that also contains the two exponential decays on day $k-1$; $\mathbf{C}$ is the output matrix that shows the influence of each state variables on the measured performance; $\mathbf{K_k}$ denotes the optimal kalman gain.

To summarise the process, KF operates at any time (any new observation). First, the basic level of performance and the last estimates of fitness and fatigue states produce the performance output through the output matrix $\mathbf{C}$. Then, the output is compared to the observed performance, and a residual error e_k is calculated. In case of no measurement, $e_k = 0$. Calculation of the Kalman gain $\mathbf{K_k}$ follows, according to the state and observation noises variances. The estimate of both fitness and fatigue states is corrected by the KF using the error feedback. Finally, a new training impulse is added to the state variables after multiplication with the time-varying input matrix $\mathbf{B_k}$.

An application of the proposed DR extension DR_{kalman} was made by authors using data from five swimmers, collected over 160 days and defining *training input* as intensity-weighted training volume, according to (Mujika, Busso, Lacoste, et al. 1996). Performances were daily collected and expressed as the mean velocity reached for 60m. When the model was trained (*i.e. fitted*) over the first half of the data, used for predicting the second half, error of prediction reported by DR_{kalman} was significantly lower for four of the five swimmers (mean $MAPE = 4.12\%$ and $MAPE = 3.56\%$ for DR and DR_{kalman}, respectively). Lower gain term values were found for k_h (i.e. the fatigue gain term, about $1e^{-06}$) than for k_g (i.e. the fitness gain term, about $1e^{-04}$), describes a lower fatigue response than the fitness one for the given data. In addition, a high variance in parameter estimates between subjects was reported, which is in line with some results observed in the literature (Hellard, Avalos, Lacoste, et al. 2006). Hence, similar performances were obtained with a model that uses generalised parameters, calibrated across all subjects (mean $MAPE = 3.57\%$ for DR_{kalman} and mean $MAPE = 5.00\%$ for DR over the test data set). It thus questions the absolute need to calibrate models per subject, particularly when the sample size available for training the model is small.

Conclusion about fitness-fatigue models.

Fitness-fatigue models are precursors in the modelling of training effects on performance and remain still used so far. Having benefited from physiology, physics and mathematics experts, several extensions of the original FFM have emerged for better describing changes in athletic performance as well as of improving the predictive capability of models. Hence, FFMs consist of linear time-invariant models and linear time-varying models, heterogeneous in terms of model complexity. In addition, using the formalism of transfer functions as a model of relationships between the training doses and fitness and fatigue states allows for finding an optimal control law (i.e. the training doses), analytically provided from the algebraic structure of the dynamical system and the expected output. Consequently, it does not require any simulation of training protocols to find the optimal training sequences, unlike some other statistical approaches presented later in the manuscript(see Section 3.2.2).
With the exception of a recent extension (Piatrikova, Willsmer, Altini, et al. 2021), FFMs remain univariate linear model which fully rely on a single input –a training load parameter– for estimating states variables. Extending models to more complex functions or using algorithms to correct predictions errors like the Kalman filter might compensate for the simplifying consideration of biological processes involved by exercise and underlying athletic performance.

On another note, few are the studies that have performed model validation procedures by training the model and making predictions on two different subsets of data (Chalencon, Pichot, Roche, et al. 2015; Kolossa, Azhar, Rasche, et al. 2017; Ludwig, Asteroth, Rasche, et al. 2019). If that might not be so prejudicial for describing changes in athletic performance, it may become a real problem when the model is intended to predict future changes in athletic performances. Furthermore, FFMs are usually considered and intended to be interpretable models (which is another good point for FFMs). While practitioners might draw helpful insights from model parameters (mainly the time constant for exponential decays) for optimising training programs, they should also have confidence in the model, particularly its ability to generalise.

1.2.2. Analogies with the fitness-fatigue framework

1.2.2.1. Training effects and performance potential

In analogy to FFMs, a metamodel for simulating the training effect on athletic performance has been provided by Perl 2001, under the name of PerPot model. PerPot is based on the super-compensation theory –like FFMs– but comes with two fundamental notions, the potential described by resources that are not infinite and that deplete in response to training stimuli, and the overload in which the system is no longer able to perform adequately and begin to degrade. Since the mathematical framework underlying PerPot would be for the benefit of any systems

as long as they control themselves using antagonistic control structures, authors defined PerPot as a metamodel (Perl 2001). A diagram of the PerPot structure is presented in Figure 1.5.

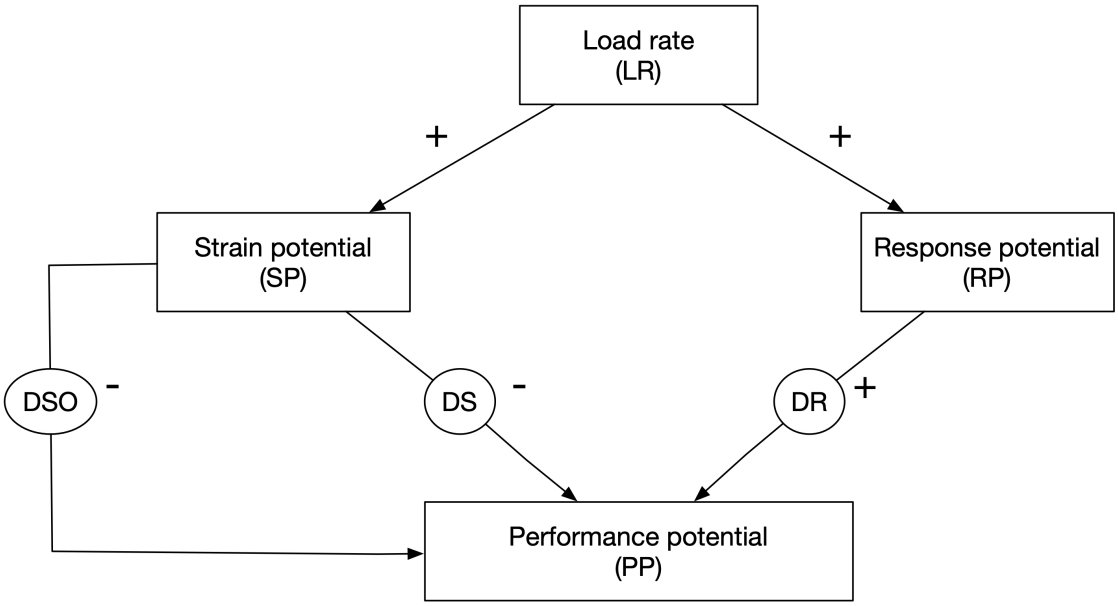


Figure 1.5. – Diagram of the PerPot model, adapted from Perl 2001. The terms DSO, DS and DR denote the delays of strain overflow, delays of strain flow and delays of response flow, respectively.

According to Figure 1.5, the model consists in a general function $w'(t)$ that gives a training load rate (LR) as model input. Each input feeds both strain potential (SP), and response potential (RP) reserves that affect performance in an antagonistic way through delayed strain (DS) and response (DR) flows, analogously with the fitness and fatigue states from FFMs. In the case of SP saturation, such as following intense and repeated training sessions over a short period, an overflow of strain potential occurs and enhances a strain overflow delay (DSO) that negatively affects the potential of performance (i.e. decreases). In an extreme situation, a severe overtraining would be characterised by empty reserves of performance potential (PP) in response to a heavy strain overflow.
Hence, the performance potential is driven by the two strain and response potential states and varies according to the positive and negative flow delays. Suppose we assume that $DS = DR$, there is no effect of training on performance potential. Otherwise, if $DS > DR$ we can expect an increase in PP and conversely, if $DS < DR$ the model predicts a decrease in performance.

According to Perl 2000, let us define the main equations of the model :

$$\text{raising potentials} = \begin{cases} SP & := SP + LR \\ RP & := RP + LR \end{cases}$$

$$\text{computing rates} = \begin{cases} SR & := \frac{1}{DS}\, min\big(min(1, SP)\,,\, max(0, PP)\big) \\ RR & := \frac{1}{DR}\, min\big(min(1, RP)\,,\, min(1, 1 - PP)\big) \\ OR & := \frac{1}{DSO}\, max(0, SP - 1) \end{cases}$$

$$\text{updating potentials} = \begin{cases} SP & := SP - SR - OR \\ RP & := RP - RR \\ PP & := PP + RR - SR - OR\,. \end{cases}$$

Here, LR denotes the load rate; SR, RR and OR are the strain, response and overflow rates. All upper limits and lower limits are normalised to 1, 0, respectively. PerPot model parameters are usually estimated from the pool of observations by minimising the RSS between predicted and observed performances.

To our knowledge, PerPot remains scarcely used for modelling the effect of training on athletic performance (Ganter, Witte, and Edelmann-Nusser 2006; Pfeiffer 2008; Torrents, Balagué, Perl, et al. 2007). Pfeiffer 2008 has compared the two antagonistic concepts FFM and PerPot presented so far. The author found that using data from six female gymnasts, the performances reached by PerPot model were greater than ones from the original FFM (mean relative deviation were equal to 2.48% and 2.78% for PerPot model and FFM, respectively). In addition, PerPot model seemed to better fit real performances than FFM, supporting a more adaptive internal model dynamics of the model (Pfeiffer 2008).

In conclusion, the PerPot model seems conceptually rich since it accounts for the collapse effect of an overloading period, the fall in physical abilities following a de-training period and the long-term behaviour of the training - performance relationship (Hellard, Avalos, Lacoste, et al. 2006). In addition, setting limit values for states variables enables the model to predict changes in performance in a non-linear fashion. According to some basic physiological processes and, more generally, normal overtraining and de-training processes occurring during training, PerPot model seems more suitable than fully linear time-invariant models such as FFM. However, applications that use PerPot model remain few and mainly based on simulated data Perl 2001. Further applications using real data –even historical– would be welcomed.

1.2.2.2. Indirect pharmacodynamic models for athletic performance

In pharmacodynamics, indirect response models have been developed for predicting the pharmacologic behaviour of drugs through the time course of pharmacodynamic responses with plasma drug concentrations (Sharma and Jusko 1998). Such models can describe rebound phenomena and drug tolerance in response to the accumulation of drug administration (Sharma, Ebling, and Jusko 1998; Sharma and Jusko 1998). Since responses to many drug administration are indirect (e.g. anticoagulant effect of warfarin, receptor/gene-mediated induction of the enzyme tyrosine aminotransferase) (Dayneka, Garg, and Jusko 1993), an analogy with changes in athletic performance in response to exercise can be made since both rebound phenomena and *exercise* tolerance exist. Indeed, physical exercise involves acute responses that are a primary stimulus for training-induced adaptations, activating a secondary signal that translates the primary stimulus into physiological adaptations. As an example, physical exercise induces cellular disruptions that activate various signalling pathways (i.e. secondary signals), involved in physiological functions such as tissue remodelling (D'Antona, Lanfranconi, Pellegrino, et al. 2006), phenotypic plasticity (Coffey and Hawley 2007) and more generally organismal growth and homeostasis functions (Laplante and Sabatini 2012; Song, Moore, Hodson, et al. 2017). Hence, changes in athletic performance depend on the rate of secondary signal activation that persists after the cessation of training and, therefore, the accumulation of training sessions over a finite temporal horizon and their magnitude (by analogy to exercise intensity).

From this basis, using such an indirect secondary signal in the modelling offers an alternative to FFMs, that allows for describing non-linear training effects on performance in a pharmacodynamical fashion (Busso 2017). In terms of adaptations to exercise, it might consider that acute fatigue negatively impacts performance over a short time frame and a decrease of adaptation (i.e. tolerance) or increase of maladaptation due to excessive repeated training loads. Also, such models may address some of the theoretical drawbacks ascribed to FFMs and stated in Sections 1.2.1 and 1.2.1.1.

According to Busso 2017, we provide the basic equations that lead to the formulation of four secondary signal models, which vary in terms of complexity. First, let consider changes in physical performance resulting from the sum of the accumulated responses to each training session produced by an indirect response. Accordingly, change in performance y without any training is defined such as

$$\frac{dy}{dt} = k_{on} - k_{off} \cdot y \,.$$

Here, k_{on} denotes a zero-order rate variable (i.e. linear) for the production of performance and k_{off} a first-order rate constant (i.e. exponential) for its removal. Stimulation of performance production at time t depends on a training load W, and a secondary signal of the same amount that decays with a first-order rate constant

k^s_{out}. Meanwhile, the signal is transformed into performance with a first-order rate constant k^s_{on} and added to a baseline constant k^0_{out}. Hence, we write the rate of change in signal S such as

$$\frac{dS}{dt} = W - k^s_{on} \cdot y \,.$$

Hence, the general formulation for the production of performance P becomes

$$P = k^0_{on} + k^s_{on} \cdot S \,.$$

All models proposed by Busso 2017 are presented in Figure 1.6. First, model T is the simplest formulation for indirect pharmacodynamic response where changes in performance depend on the secondary signal that counteracts the natural dissipation of performance. A second model TI includes inhibition function that inhibits the production of performance by the secondary signal. Its value for a given day n might be defined as

$$I_n = k^i_{in} \cdot w_n \,,$$

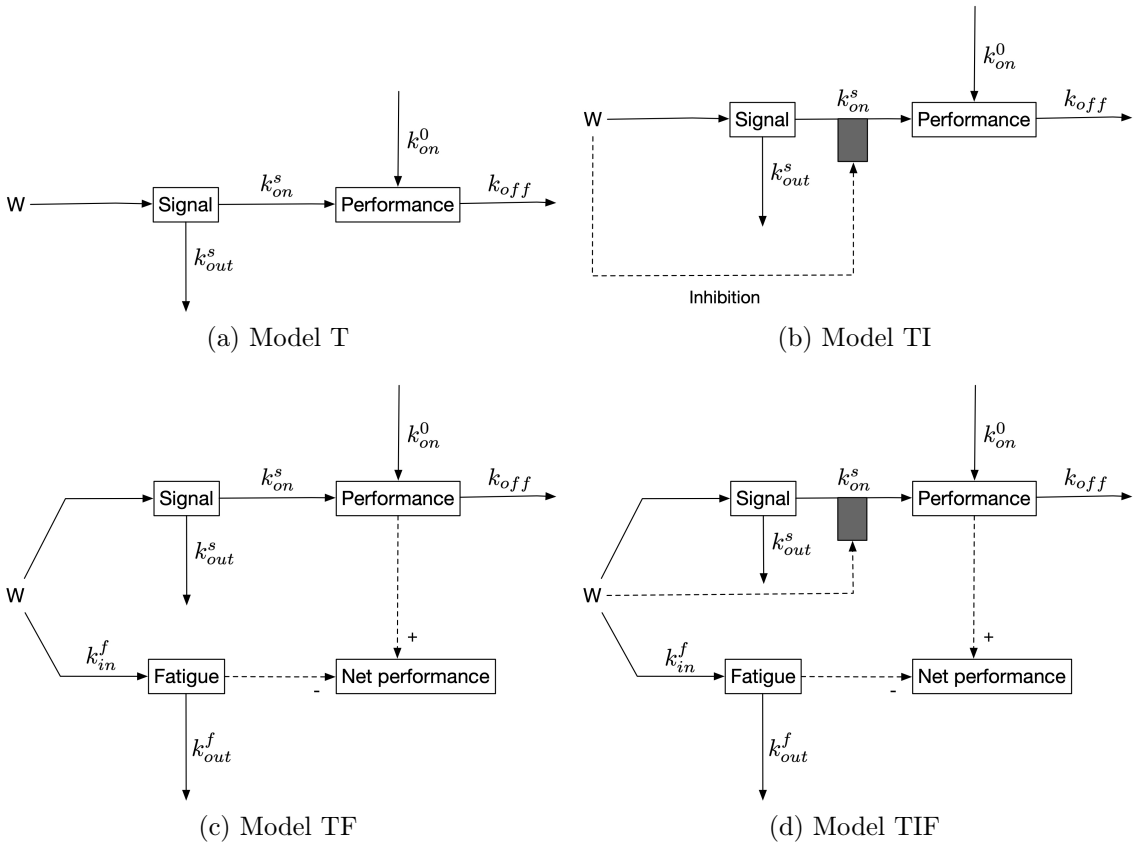
with k^i_{in} being a constant. Hence, the general formulation of performance production P becomes

$$P = k^0_{on} + k^s_{on} \cdot S \cdot (1 - I) \,.$$

By analogy with FFMs, two more models TF and TIF are based on model T and TI, respectively, and include a fatigue component F that negatively impacts the performance. The net performance becomes the difference between performance and the fatigue estimate. The production of fatigue is proportionally related to training loads, weighted by a constant k^f_{in} that decays according to a first-order rate constant k^f_{out}. The rate of change in fatigue is

$$\frac{dF}{dt} = k^f_{in} \cdot W - k^f_{out} \cdot F \,.$$

The discretisation of model equation is available in Appendix D



(a) Model T (b) Model TI

(c) Model TF (d) Model TIF

Figure 1.6. – Block diagrams of four secondary signal models, reprinted from Busso 2017.

As a result, Busso 2017 found that using indirect pharmacodynamic models with secondary signal are valuable for describing changes in athletic performance. Using data from a previous study (Busso 2003), author found that model TI was the most performing among the four propositions, reaching similar R^2 values than using a DR model ($adjR^2_{TI} = 0.945 \pm 0.019$ and $adjR^2_{DR} = 0.944 \pm 0.042$ from a previous study) (Busso 2003). In addition, adding a fatigue component to the model did not significantly improve the model performances, except for one subject. That is in line with literature since the most *complex* models (assuming that *complexity* is related to the number of components included in the model) did not ensure the best performances (Busso, Candau, and Lacour 1994). Once again, one should consider the population studied in studies before generalising in so far as athletes may train with much higher workloads than those studied here, which might render a fatigue component valuable in such cases.

In conclusion, the models presented by Busso 2017 appear to be valuable non-linear alternatives to classical FFMs. Using inhibition function addresses the main limitations of the original FFM by lowering or negatively impacting the performance gains through the use of inhibition term as a function of past training sessions.

Yet, such models' ability to describe and predict athletic performance and find an optimal training sequence remains to be further investigated in other populations and sports.

1.2.3. Contributions from statistics and computer science

In Sections 1.2.1 and 1.2.2 we presented some mathematical modelling of athletic performance using models based on first-order transfer functions. These models have the particularity of being highly interpretable and benefit from their algebraic structure for figuring out an optimal training dose. In addition, models are based on expert hypotheses raised by years of biological, physiological more generally exercise science research, making them of great rich. Nevertheless, their main limitation comes with their univariate consideration of changes in athletic performance, since they use information from a single measurement as model input, independently of the number of produced features from that measurement (i.e. usually a training load index, see 1.1 for details).

From a statistical point of view, models presented so far are linear models (time-varying or time-invariant), a basic framework in statistics. In sports sciences and mainly when dealing with athletic performance issues, some more elaborated statistical approaches came with a different approach to the problem. Indeed, the rise of wearable sensors and the variety of training load quantification methods create a pool of variables into which probably useful information are located. Thus, it is a matter of finding pertinent information in the midst of plenty of data and finding an optimal function that maps a set of explanatory variables to the target rather than dealing with a unique independent variable. By definition, models that would account for a vast data source are multivariate models of an infinite complexity that simultaneously brings more statistical power and potential drawbacks.

1.2.3.1. Statistics and multivariate linear models

Firstly, simplest cases for modelling the relationship between training and performance remain regression tasks using multivariate linear models. Briefly, we present the basic formulation of a multiple linear regression. Let consider a sample $\{y_i, x_{i,1}, \ldots, x_{i,p}\}_{i=1}^{n}$ of n observations with y a dependant or response variable and a vector $\mathbf{x}$ of j independent variables, and a random error term ϵ. We can model the relationship between y_i and the predictors as defined in the sequel :

$$y_i = \beta_0 + \beta_1 x_{i,1} + \ldots + \beta_p x_{i,p} + \epsilon_i \,, \quad i = 1, \ldots, n \,. \tag{1.41}$$

In analogy with FFMs presented in Section 1.2.1, the model can be re-written as

$$y_i = \beta_0 + \beta_1 \Phi_1 x_{i,1} + \ldots + \beta_p \Phi_p x_{i,p} + \epsilon_i \,, \quad i = 1, \ldots, n$$

where $\Phi_1, \ldots, \Phi_p$ represent either a linear or a nonlinear function, $\beta_1, \ldots, \beta_p$ denote the estimated coefficient for each independent variable.

To date, a few authors have attempted to model the relationship between training and performance using multivariate linear models (LM) or extensions (Avalos, Hellard, and Chatard 2003; Hellard, Avalos, Hausswirth, et al. 2013; Hellard, Scordia, Avalos, et al. 2017; Knobbe, Orie, Hofman, et al. 2017; Mujika, Chatard, Busso, et al. n.d.). According to Avalos, Hellard, and Chatard 2003, using multivariate models allow for modelling the relationship between training loads and changes in athletic performance, accounting for other training-related variables such as training periods (e.g. short, mid and long-term training periods). In that case, it provides a different approach than the one from FFMs in the way of accounting for the effect of time on performance outcomes, where each period influences the response linearly. Other variables of importance such as technical level, environmental factors and psychological states can also be included in the model, known to be of importance for athletic performance modelling purposes (Hellard, Avalos, Lacoste, et al. 2006).

Since a singularity in training responses is largely assumed in the literature (Hellard, Avalos, Lacoste, et al. 2006; Mujika, Chatard, Busso, et al. n.d.), we can expect a different effect of a particular time of exercise on the body's response among athletes. One would note that it is a cornerstone of FFMs, initially intended to be fitted per athlete. That is, modelling the effects of training on athletic performance from data of a set of athletes should consider individual aspects of adaptations to training. Using LM, individual responses might be considered either, for the simple case, by setting athletes as a covariable (i.e. it weights the response linearly), or by extending the basic formulation of LM (see Equation 1.41) to more complex models.

Among extensions of LM, we note linear mixed models (LMM). Avalos, Hellard, and Chatard 2003 are the first authors that used LMM for predicting athletic performances. In brief, a LMM can be written such as

$$Y_i = X_i\beta_I + Z_i b_i + \epsilon_i \,,$$

with $b_i \sim \mathcal{N}(0, D)$, $\epsilon_i \sim \mathcal{N}(0, \sum_i)$ and b_i's independent of ϵ_i's. In the model, we consider Y_i being the performances of a subject i, $X_i\beta$ denotes the common term in the model, $Z_i b_i$ is a specific term to each subject i and ϵ_i denotes the random error term. The regression coefficients β are fixed effects applied to the set of variables X_i of subject i and b_i denotes the random effects that represent individual regression coefficients for subject i applied to Z_i, the subset of variables that are specific to each subject. Random effects b_i are supposed to be normally distributed with the covariance matrix D and ϵ_i is also supposed to be normally distributed with the autoregressive covariance matrix $\sum_i$ (Avalos, Hellard, and Chatard 2003; Verbeke 1997).

In their study, authors defined training load variables according to specific training periods as independent variables. Through principal component analysis (PCA) and cluster analysis, they detected patterns in training responses among subjects that were included as covariates in the model (Avalos, Hellard, and Chatard 2003; Hellard, Scordia, Avalos, et al. 2017). Such models highlighted the fact that athlete responses to a given training vary according to the time or training period, which is in line with precedent findings (Banister, Carter, and Zarkadas 1999; Busso, Denis, Bonnefoy, et al. 1997). Moreover, it is of value for any training programming purposes since careful attention is given to the training overload and taper cycles programming (Hellard, Avalos, Hausswirth, et al. 2013).

Using LMM in athletic performance modelling tasks showed greater model performances than the original FFM (Hellard, Avalos, Lacoste, et al. 2006) and, thus, a better comprehension of individual training responses. On the statistical side, LMM benefits from a larger sample size for inferring parameters since the model accounts for observations from all subjects without neglecting their singularity in training responses. It is of importance when we attempt to predict real athletic performances that are few over a season while providing *accurate* parameter estimates (Cui and Gong 2018; Kelley and Maxwell 2003).

Including many variables in a LM ensures to catch all the available information, but it also may raise multicollinearity problems. Multicollinearity is the occurrence of high intercorrelation between explanatory variables in a multiple regression model. It can produce wider confidence intervals and thus affects the statistical inference, leading to misleading results. There are many methods for dealing with multicollinearity issues in statistics, notably through the use of regularisation procedures (Bickel, Li, Tsybakov, et al. 2006; Friedman, Hastie, and Tibshirani 2010; Hoerl and Kennard 1970; Tibshirani 1996; Witten and Tibshirani 2009; Zou and Hastie 2005). These procedures come with the aim of stabilizing an estimator by shrinking the space of solutions around null. In practice and using a regression problem as an example, we can modify the cost function $J(\theta)$ such as

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^{m} \left(h(x_i) - y_i\right)^2 + P(\lambda, \Theta)\,.$$

In this equation, a penalty function $P(\lambda, \Theta)$ is added to the ordinary least square cost function, with $h(x_i)$ being a linear model. The severity of the penalty is given by the parameter λ.

While regularisation procedures based either on penalisation or features transformation are widely used in several domains, only a few studies used them for athletic performance modelling purposes within longitudinal follow-up contexts (Knobbe, Orie, Hofman, et al. 2017; Kosmidis and Passfield 2015). In the latter, the authors used a LASSO regression for detecting an optimal aggregation of discrete sequences

of events and then to predict performance outcomes in elite cyclists (Lotto NL-Jumbo team) (Knobbe, Orie, Hofman, et al. 2017). By doing so, the authors provided an attractive approach using both an expert hypothesis through exponential aggregations of training load measurements (like fitness-fatigue features from the original FFM) and other variables that may affect the performance outcomes.

1.2.3.2. A machine-learning approach to the problem

Statistics aim to draw population inference from data upon rigid assumptions, thus making an explicit understanding of how a system behaves. Close but slightly different, ML aims to find generalisable predictive patterns upon minimal assumptions, intended for making future predictions rather than inferring –in our case– relationships between training and athletic performance outcomes. Although statistical and ML methods may be used for both inference and prediction, ML benefits from powerful algorithms for predicting non-linear patterns and interactions (IJ 2018).

Looking back to the study of Knobbe, Orie, Hofman, et al. 2017, authors used *Subgroup Discovery* method (Atzmueller 2015) in order to deal with the assumed non-linear dependencies between training and performance. Subgroup Discovery is a supervised learning method that aims at discovering relations between a dependant and several independent variables for inductive and exploratory data analysis tasks. Hence, *Subgroup Discovery* allows for finding candidates within subgroups, according to a quality measure (a function predefined by the analyst). Hence, authors showed that using such non-linear modelling alternatives to LM provided better model performances and useful insights intended to coaches for their training programming tasks.

On another note, a few authors have attempted to predict changes in athletic performances using ML algorithms. artificial neural network (ANN) were firstly suggested to be a valuable alternatives to FFMs, as presented by Edelmann-Nusser, Hohmann, and Henneberg 2002. In their study, authors modelled the swimming performances of an elite female swimmer in the finals of the Olympic Games in Sydney. Through computation of feed-forward ANN (one input layer with ten neurons, one hidden layer with two neurons and a single output neuron), authors reported highly accurate performance of the Olympic final with an error of prediction equal to 0.05s only. In their methods, authors used data from another swimmer to fill the lack of data and pre-train the model, despite the singularity in the training responses exposed so far. Models were validated by a leave-one-out procedure that consists in dropping one sample out for model validation. Even though the results provided by Edelmann-Nusser, Hohmann, and Henneberg 2002 should be taken with caution since they predicted only one observation from a single athlete (we may expect some different results on other data and athletes), ANN remains promising for predicting athletic performances since linear multivariate models can only approximate non-linear problems. The results presented in this

study were further supported by authors that predicted swimming performances in the Olympic Games of Athens, 2004 (Pfeiffer and Hohmann 2012).

More recently, two studies attempted to model responses to training using ANN (Carrard, Kloucek, and Gojanovic 2020; Mitchell, Rattray, Fowlie, et al. 2020). In the first, the authors focused on modelling adaptations and maladaptations to training using geometric optimisation. Combinations of weekly *external* and *internal* training loads averages formed the model inputs, according to the recommendations of Bourdon, Cardinale, Murray, et al. 2017 (see Section 1.1 for details). The output was described as a percentage of the personal best time (%PBT) in which each new value of %PBT represented either an adaptation (exceeding %PBT baseline) or a maladaptation (below %PBT baseline). Although the results showed great model performances (an average goodness-of-fit of 95%) and provided clues to coaches, they cannot be considered valid for future modelling tasks since authors did not perform any model validation due to insufficient sample size. Therefore, such an approach focusing on predicting adaptations to training rather than a raw athletic performance remains to be further investigated and might valuable insights to coaches. In the latter, authors compared the performances of FFMs (original and DR FFMs, see Section 1.2.1.1 and 1.2.1.4 for mathematical formulation) (Banister and Hamilton 1985; Busso 2003), exponentially weighted moving averages and ANN to a rolling averages model and according to two TL quantification methods (a five-zone and a seven-zone quantification methods), using data from three elite swimmers. Neural networks in which training load aggregates differ (short, mid and long term periods according to Avalos, Hellard, and Chatard 2003 and four exponential increasing time frames) were cross-validated along with other compared models, using a leave-one-out cross-validation (LOOCV) on a training data set (80% of the full data set). The results reported in this study showed that all models were more accurate than the rolling average model, with the lowest errors attributed to ANN. In addition, a TL quantification based on five zones seemed to be more efficient, although a seven-zone should bring more information regarding the changes in exercise intensity.

The use of ANN goes beyond the longitudinal modelling of athletic performance and find various applications in sport sciences (e.g. injury prevention, tactical and technical analysis) (McCullagh and Whitfort 2013; McCullagh et al. 2010; Perl, Tilp, Baca, et al. 2013; Pfeiffer and Hohmann 2012; Pfeiffer and Perl 2006; Ruddy, Shield, Maniar, et al. 2018; Schneider, Hanakam, Wiewelhove, et al. 2018; Silva, Costa, Oliveira, et al. 2007). Despite being powerful and flexible models, they do not represent unique *world class* models to be used since many ML models (e.g. decision trees, random forest, support vector machines for supervised learning) might be valuable for athletic performance modelling purposes (Wolpert 2021; Wolpert and Macready 1997). That being said, choosing a particular class of functions should be driven by data properties, assumptions and hypotheses (if any), and the objective behind the modelling.

1.3. Conclusion

In this chapter, we defined training load and presented the main quantification methods based on physiological principles. Be the training load measures come from either objective or subjective measures; we should consider them together for athlete monitoring purposes. When it comes to model the relationship between the effect of training and athletic performance or injury occurrence, plenty of models that come from more or less various mathematical branches (*i.e.* control theory, statistics, ML) can be used. Even though some machine-learning approaches seem to differentiate from other approaches with a greater pertinence, there is still no consensus on the class of models to be the most valuable for athletic performance modelling according to the *no free lunch theorem* (Wolpert and Macready 1997). Furthermore, combining physiological, statistical, and computer science knowledge opens the gate for a better comprehension of athlete response to training to optimise training protocols in the best way.

2. Training loads: from an objective training-load quantification to physiological responses

Table of contents

2.1 Resistance training: from an objective training-load quantification to physiological responses . . . 88
2.1.1 Introduction . . . 88
2.1.2 Experimental setup . . . 89
2.1.2.1 Population studied . . . 89
2.1.2.2 Modelling Torque – velocity profiles . . . 89
2.1.2.3 Resistance exercise protocols . . . 92
2.1.3 Data collection . . . 94
2.1.3.1 Systemic measurements . . . 94
2.1.3.2 Local measurements . . . 96
2.1.4 Physiological responses to resistance exercises . . . 101
2.1.4.1 Metabolic and hormonal responses to resistance exercises . . . 101
2.1.4.2 Cardiac kinetics and pulmonary gas exchange in response to resistance exercises . . . 104
2.1.4.3 Neuromuscular responses to resistance exercises . . . 112
2.1.4.4 Conclusion . . . 125
2.2 Towards a new model of training effect quantification . . . 125
2.2.1 A physiological model of training load quantification in resistance training . . . 126
2.2.2 A linear combination of quantification methods and exercise related variables . . . 128
2.2.3 Relationship between training load quantification methods and physiological responses . . . 133
2.2.4 Conclusion . . . 134
2.2.5 Take-home message . . . 136

2.1. Resistance training: from an objective training-load quantification to physiological responses

2.1.1. Introduction

Popular among amateurs and athletes aiming for a performance improvement (Deschenes and Kraemer 2002; Hamilton, Paton, and Hopkins 2006), injury prevention (Faigenbaum and Myer 2010) and health (Feigenbaum and Pollock 1999; Williams, Haskell, Ades, et al. 2007), RT induces a wide range of adaptations located at physiological (Deschenes and Kraemer 2002; Fry 2004; Kraemer, Deschenes, and Fleck 1988), hormonal (Kraemer and Ratamess 2005; Walker, Häkkinen, Newton, et al. 2020), neuromuscular (Häkkinen, Alen, Kraemer, et al. 2003) and cardiovascular levels (Fleck 1988). By considering a simple dose-response model in which training loads (i.e. the "dose" being a quantitative representation the work produced or the stress applied to the body) induce adaptations (i.e. the "response", illustrated by either such aforementioned adaptations detailed in Section 1.1.1.2 for details), monitoring RT forms the basis of any training programming optimisation. Hence, inappropriate training doses –commonly perceived as unbalances between "positive" and "negative" adaptations– may lead to performance impairments and injuries (Fry and Kraemer 1997; Kibler, Chandler, and Stracener 1992).

Monitoring RT is essential in order to understand the effects of RT on the underpinning adaptations and changes in athletic performance following training cycles. In section 1.1.1.2, we presented the most common methods for quantifying TL in RT. Summing up, using subjective approaches through RPE and sRPE measurement, as well as more objective approaches based on volume load indexes, remain gold standard practices (Haff 2010; Helms, Kwan, Sousa, et al. 2020). Behind these quantitative measures, monitoring aims at bringing to light the stress induced by the training sessions in a simple, practical, valid and reproducible way. However, the methods presented so far showed significant limits and remain inaccurate in their physiological meanings so far (Genner and Weston 2014; Marston, Peiffer, Newton, et al. 2017).

Understanding physiological responses to exercise and, therefore, the changes in athletic performance (if we assume that they are correlated –even partly– to physiological responses and adaptations) implies building a model of these relationships. Some authors have shown that skeletal muscle adaptations to RT were linearly correlated with exponential growth transfer functions in rodents (Philippe, Borrani, Sanchez, et al. 2019). Hence, before any modelling of athletic performance, efforts could be targeted at providing valuable inputs (e.g. training load indexes) supported by observed physiological mechanisms.

In the following sections, we investigate the physiological responses to resis-

tance exercises in controlled conditions. Then we highlight relationships between well-known methods used for quantifying TL, namely VL and RPE, and acute physiological responses to exercise. We also provide a new method of TL quantification based on training stress index (TSIdx) supported by physiological responses to RE and intended to be used for monitoring purposes within a systemic approach.

2.1.2. Experimental setup

First, individual T-V profiles were modelled according to leg extension tasks performed on an isokinetic dynamometer. Then, the validity of each TL quantification method was investigated over three volume-equated protocols, individually calibrated according to T-V profiles. These three sessions were performed in randomised order and are described in Section 2.1.2.3.

An overview of the setup is represented in Figure 2.6.

2.1.2.1. Population studied

Fifteen trained adults (eleven male age: 26.9 ± 3.45 years, mass: 78.11 ± 7.56 kg, body fat: 11.43 ± 3.39 % body mass and four female age: 21.75 ± 1.54 years, mass: 60.63 ± 3.92 kg, body fat: 21.10 ± 3.39 % body mass) voluntary participated in the study. The inclusion criteria for participation were (i) having experienced resistance training continuously for at least six months at the onset of the study, (ii) being familiar with resistance exercises performed at maximal intensities and (iii) not having any current, recurring injury on lower limbs or functional limitations regarding a knee extension task performed at maximal intensity. In addition, participants were asked to respect their usual diet all along the study period. The testing session was performed on different days but in respect of the circadian rhythm.

2.1.2.2. Modelling Torque – velocity profiles

A first testing session was performed in order to model the individual T-V of a knee extension task. First, the participant performed a cardiovascular warm-up on ergo-cycle (Ergoselect, ergoline GmbH, Germany) that consisted of 4 minutes cycling at 50W and a cadence of 50 to 60 revolutions per minute.

Then, the participant was seated on an isokinetic dynamometer (Biodex system 3, Biodex Medical Systems, USA). The shaft was aligned with the axis of rotation of the knee joint to be tested. Torso, waist, pelvis and working limb were secured with straps. Handles were disposed on both sides of the chair, on which open hands are placed during exercise. A shin pad attached to the distal extremity of the mechanical arm was firmly secured to the working leg about 5 cm above the medial malleolus. Once the participant was poised, lever arm amplitudes were recorded in internal to external positions (i.e. from naturally bended knee to fully extended

knee). The working leg was weighed in an external position and considered in the isokinetic measures.

A specific warm-up followed the setting step. Participants were asked for performing four repetitions of concentric extension at $1.047\,\mathrm{rad\,s^{-1}}$ with a self and progressive increase of intensity plus two repetitions at maximal intensity. The knee extension being the only movement of interest, knee flexion was assisted by returning to the initial position with a velocity of $5.236\,\mathrm{rad\,s^{-1}}$ producing no resistance.

After a passive rest period of four minutes, participants performed seven series of concentric extensions 3 minutes apart at the following velocities in a quasi-randomised order: $0.524\mathrm{rads}^{-1}$, $1.047\mathrm{rads}^{-1}$, $1.570\mathrm{rads}^{-1}$, $2.094\mathrm{rads}^{-1}$, $2.618\mathrm{rads}^{-1}$, $3.142\,\mathrm{rad\,s^{-1}}$, $3.665\,\mathrm{rad\,s^{-1}}$. In order to limit the fatiguing effect of the lowest velocities, only two repetitions were performed at $0.524\,\mathrm{rad\,s^{-1}}$ and $1.047\,\mathrm{rad\,s^{-1}}$, against three repetitions at other velocities. These velocities were performed before the sixth of the seven series. A 1-s break was set between two consecutive contractions to avoid any possible influence of stretch-shortening cycle. Using seven points enabled to model a valid and reproducible T-V profile (James, Sacco, Hurley, et al. 1994; Lemaire, Ripamonti, Ritz, et al. 2014).

As shown in Figure 2.1, T-V profiles were modelled according to an exponential function

$$f(x) = \alpha e^{kx} + c\,, \tag{2.1}$$

where α denotes a scaling factor, k denotes a negative rate constant and c is an additional intercept term, k estimates greatly decrease across the four T-V models ($k = -0.0168$, $k = -0.00989$, and $k = -0.00247$ for models in Figure 2.1a, 2.1b, 2.1c, respectively), from the most *exponential* to the most *linear* trend profile.

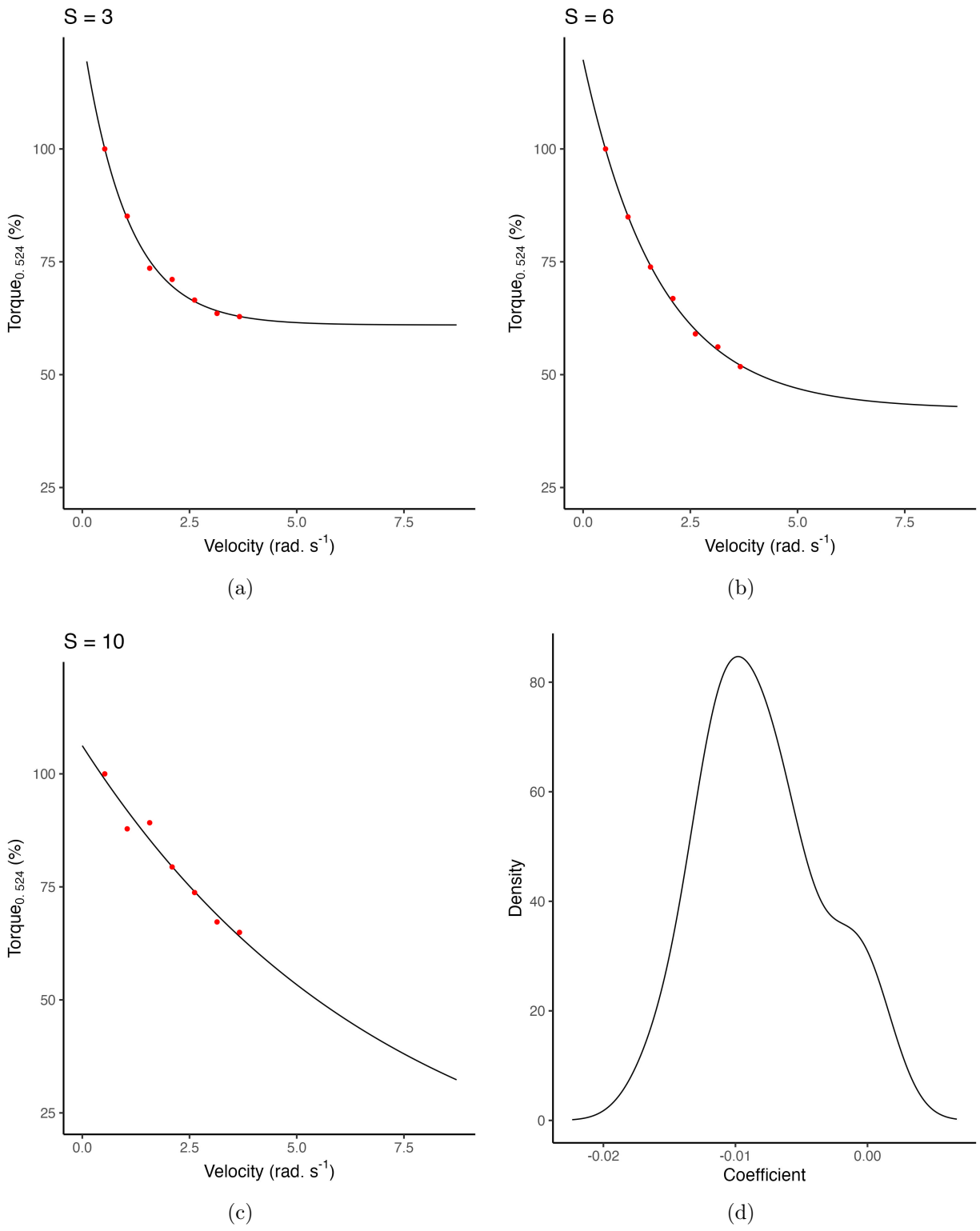


Figure 2.1. – Figures 2.1a, 2.1b and 2.1c represent T-V profiles modelled from an isokinetic knee extensions at different velocities. Values are normalised to the maximal torque measured at 0.524 rad s^{-1}. Red dots represent the measured torque for each velocity, the line represents the modelled T-V profile according to Equation 2.1. Figure 2.1d shows the distribution of decay rate constants (unit of normalised torque per rad s^{-1}, according to Equation 2.1).

Torque-velocity profiles might be fitted either with an exponential function or with a second-order polynomial function (Chow 1999). In our study, the heterogeneity of modelled T-V profiles across participants (see Figure 2.1) is in line with the literature, since authors reported either linear trends (Lemaire, Ripamonti, Ritz, et al. 2014) for profiles modelled at velocities over $1\,\mathrm{rad\,s^{-1}}$ or more likely exponential trends if a wider range of velocities is performed (Chow 1999; James, Sacco, Hurley, et al. 1994). However, differences in model shape between participants might reflect characteristics of the population since various people profiles (e.g. endurance or explosive, amateurs or professional athletes) were involved in the study.

2.1.2.3. Resistance exercise protocols

In order to assign an equated volume between low intensity (LI), moderate intensity (MI), and High intensity (HI) testing sessions (also named $C1$, $C2$, and $C3$, respectively), the equivalent relative intensity was obtained from individual T-V profiles. Then, RM and their corresponding term in relative intensity were estimated from a non-linear equation of Reynolds, Gordon, and Robergs 2006, defined in the sequel and represented in Figure 2.2.

$$y = 55.51\, e^{-0.0723x} + 48.47\,, \tag{2.2}$$

where y denotes the percentage of relative intensity (% maximal torque) and x denotes the number of expected RM.

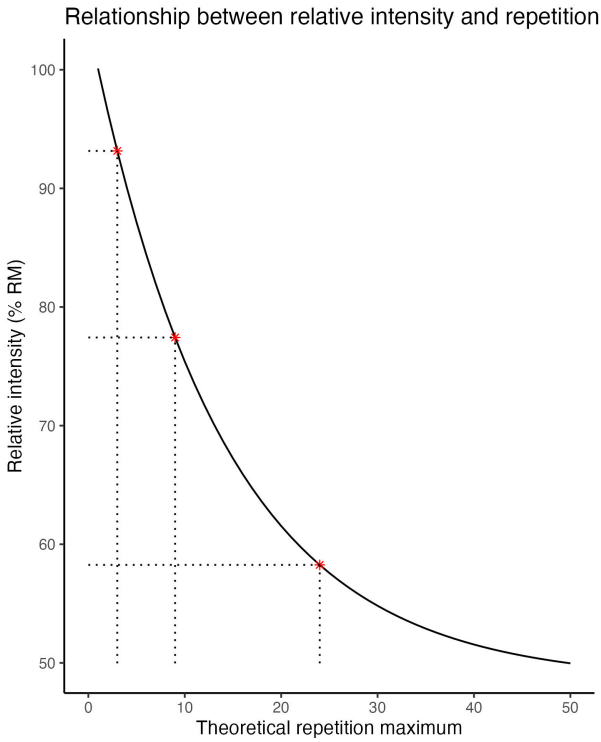


Figure 2.2. – Model of relationship between relative intensity in percentage of RM and the theoretical number of repetitions that can be achieved, according to Reynolds, Gordon, and Robergs 2006 and Equation 2.2. Red asterisks represent the retained values for the relative intensity and the number of repetition included in the protocol.

In their study (Reynolds, Gordon, and Robergs 2006), authors provided two exponential models of 1 RM prediction for leg press and bench press exercises, respectively, fitted over a sample of 70 participants. Since our exercise differs from ones of Reynolds, Gordon, and Robergs 2006, we chose to reuse the model fitted on bench press that is less conservative (less prone to over-estimation) but as accurate as of the one from leg press exercise. Hence, using Equation 2.2, the three conditions were performed at 58, 77 and 93 % theoretical maximum voluntary contraction (MVC) (i.e. the theoretical torque value for which the velocity is null), corresponding to 24, 9 and 3 theoretical RM. An overview of the testing protocols is given in Table 2.1.

Table 2.1. – Configuration of the three knee extension testing sessions

Session	Sets	Repetitions	Intensity (% MVC)	passive recovery (s)
1	1	24	58 %	N/A
2	2	9	77 %	180
3	5	3	93 %	240

2.1.3. Data collection

In order to match recordings on a single time frame, mechanical, cardiovascular and neuromuscular measurement systems were coupled using analogue signals.

2.1.3.1. Systemic measurements

Cardiac measurements

Participants wore two Electrocardiogram (ECG) sensors (Trigno EKG Biofeedback, Delsys,MA, USA) for a continuous measure of heart rate (HR) activity. Prior to starting the experiment, the quality of HR activity recording was visually checked over Q, R, S signals displayed in real-time on software EMGworks (Delsys, MA, USA). HR was further extracted from RR intervals. The continuous signal was then averaged using 10-sec bins moving average filter. Rate decay of HR during recovery was computed using a mono-exponential function. In its general form, it is defined as

$$f(x) = be^{\alpha x} + c\,, \tag{2.3}$$

with b and c two intercept terms and α a constant for exponential growth or decay.

Pulmonary gas exchange measurements

Breath-by-breath gas exchanges were analysed through a portable metabolic cart (k4b2, Cosmed, Italy), previously validated by several independent authors in locomotor activities (Brisswalter and Tartaruga 2014; Doyon, Perrey, Abe, et al. 2001; Duffield, Dawson, Pinnington, et al. 2004). Before each session, the portable system was powered on to warm up for 10 min. Calibration of the O2 and carbon dioxide (CO2) analysers was performed before every test using two-point calibration with two precision-analyzed gas mixtures (room air and a high-precision certified calibration tank gas containing O2, CO2 and balanced nitrogen). Turbine flow calibration was determined using a high-precision 3-L calibration syringe in a height-pump series. For subsequent numerical analysis, the recorded breath-by-breath gas exchange measurements were linearly interpolated on a second-by-second basis. A moving average filter was applied to the raw data in order to get an exploitable signal, since :

1. RE performed at maximal intensities involved an alteration of the respiratory frequency interspersed by short moments of apnoea.
2. RE and mainly local and mono-articular exercises usually induce a lower $\dot{V}O_2$ response than whole-body aerobic exercises, thus lowering the signal-noise ratio.

An example of the exercise identification along with recovery phase over sets is presented in Figure 2.3. From the net $\dot{V}O_2$ (i.e. the exercise $\dot{V}O_2$ minus the

resting $\dot{V}O_2$), we estimated the EE according to a caloric equivalent of $5.05\,kcal.L^{-1}$ (Wilmore, Parr, Ward, et al. 1978).

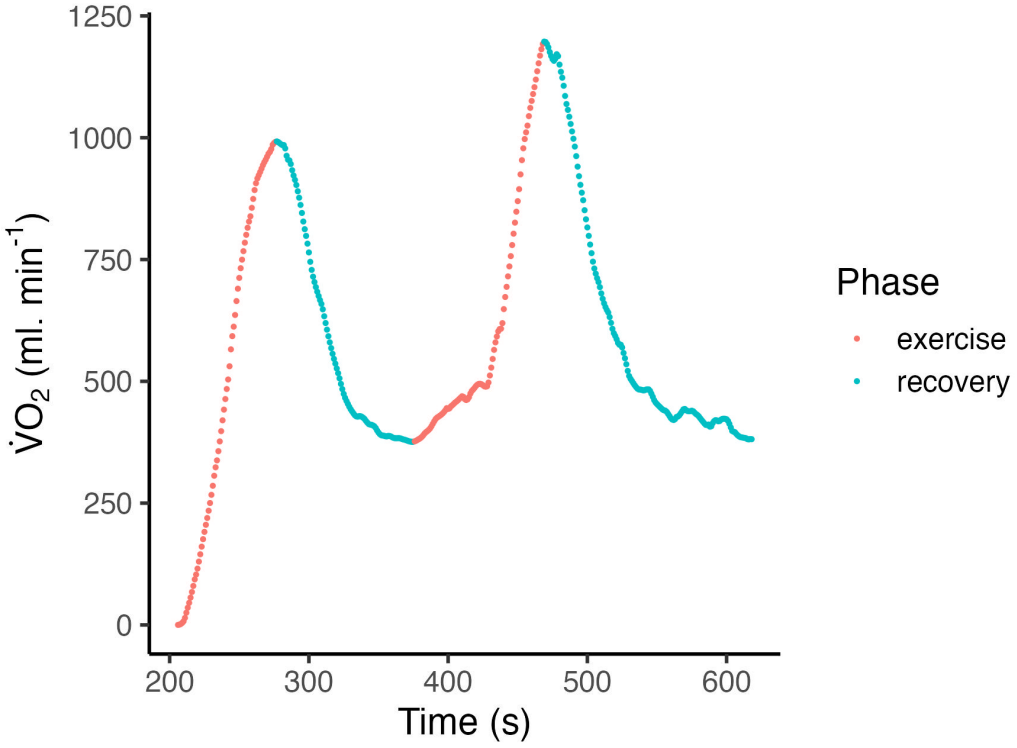


Figure 2.3. – $\dot{V}O_2$ recorded from a participant during $C2$.

Metabolic and hormonal measurements

The $[lact_b]$ were collected four times within the testing session using a finger prick before using a handheld lactate analyser (Lactate Pro, KDK Corporation, Arkray, Japan). A first sample was collected after the participant being fully equipped prior to any exercise. A second sample was taken at the onset of the testing (both global and specific warm-ups being completed). Changes in $[lact_b]$ were evaluated through 1 min and 3 min post-exercise samples in order to cover several kinetics of $[lact_b]$ responses to exercise. The Lactate Pro analyser has been widely used in sport science applications and validated as well (Baldari, Bonavolontà, Emerenziani, et al. 2009; Pyne, Boston, Martin, et al. 2000).

In addition, 10 ml of blood was collected at fingertips for $[cort_p]$ analysis. Immediately after collection, samples were centrifuged for 10 min at 2000 rev.min^{-}1. Then, plasma was collected from the centrifuged sample and stored at $-80c$. Plasma cortisol analysis was performed using Enzyme linked immunosorbent assays (ELISA) kits (Cortisol ELISA, MN, USA).

2.1.3.2. Local measurements

Mechanical measurements

For any exercises, torque (Nm), angular velocity ($rad.s^{-1}$) and position (rad) were recorded at a 148 Hz sampling frequency. From torque production over time, we estimated RFD_{0-100}, RFD_{peak} and impulsion for each repetition. Extensively investigated since many years, RFD usually reflects a neuromuscular performance (Häkkinen, Alen, Kraemer, et al. 2003; Maffiuletti, Aagaard, Blazevich, et al. 2016; Ratamess, Alvar, Evetoch, et al. 2009; Rodrı+guez-Rosell, Pareja-Blanco, Aagaard, et al. 2018) and might also be used as a surrogate measure of neuromuscular fatigue (D'Emanuele, Maffiuletti, Tarperi, et al. 2021) and muscle damages (Peñailillo, Blazevich, Numazawa, et al. 2015). Among the possible aggregates of RFD, early RFD (i.e. until 100 ms from the onset of exercise) and RFD_{peak} remain the most pertinent features for assessing neuromuscular impairments through muscle fatigue (Andersen, Andersen, Zebis, et al. 2010; D'Emanuele, Maffiuletti, Tarperi, et al. 2021).

Skeletal muscle microvascular and oxidative function measurements

Locally, skeletal muscle oxidative capacity of the VLat was evaluated by *in vivo* near-infrared spectroscopy (NIRS). Gaining popularity since the early 2000's for sports applications (Perrey and Ferrari 2018), the use of NIRS has been considered as a valid method for evaluating skeletal muscle oxygenation and oxidative metabolism (Barstow 2019; Hamaoka and McCully 2019; Nagasawa, Hamaoka, Sako, et al. 2003; Ryan, Southern, Reynolds, et al. 2013; Sako, Hamaoka, Higuchi, et al. 2001). The portable NIRS apparatus (PorLite, Artinis Medical Systems BV, The Netherlands) used in this study was a two-wavelength continuous system, which simultaneously uses the modified Beer-Lambert and spatially resolved spectroscopy (SRS) methods. Myoglobin changes were assumed to be minor compared to haemoglobin. change in tissue oxyhemoglobin concentration ($\Delta\,[0_2\mathrm{Hb}]$), change in deoxyhemoglobin concentration ($\Delta\,[\mathrm{HHb}]$) and $\Delta\,[\mathrm{tHb}]$ were measured using the difference in absorption characteristics of light at 750 and 850 nm. The TSI was calculated using SRS methods. Skinfold measurement on the NIRS location was done prior to the first session in order to ensure valid measurements regarding the adipose tissue thickness. Therefore, we can determine $\Delta\,[\mathrm{Hb_{diff}}]$ for subsequent analysis.

From $\Delta\,[\mathrm{Hb_{diff}}]$ measurements (see Figure 2.4), we estimated muscular oxygen uptake ($ml.min^{-1}$) ($m\dot{V}O_2$) through the rate decay of $\Delta\,[\mathrm{Hb_{diff}}]$ during the most representative of the first repetitions per series for which the ischemia arterial occlusion remains unaltered (Ferrari, Muthalib, and Quaresima 2011). The method is shown in Figure 2.12. In addition, we computed the rate of PCr re-synthesis using a mono-exponential model fitted on $\Delta\,[\mathrm{Hb_{diff}}]$ kinetics during recovery, such as

$$y(t) = b - Ae^{-1k}$$

where $y(t)$ denotes the relative rate of muscle oxygenation at time t, b is the rate of muscle oxygenation at the end of exercise, A denotes the difference between rates of resting muscle oxygenation and b. Finally, k being the rate constant such as an index of muscle mitochondrial capacity, according to Hamaoka and McCully 2019.

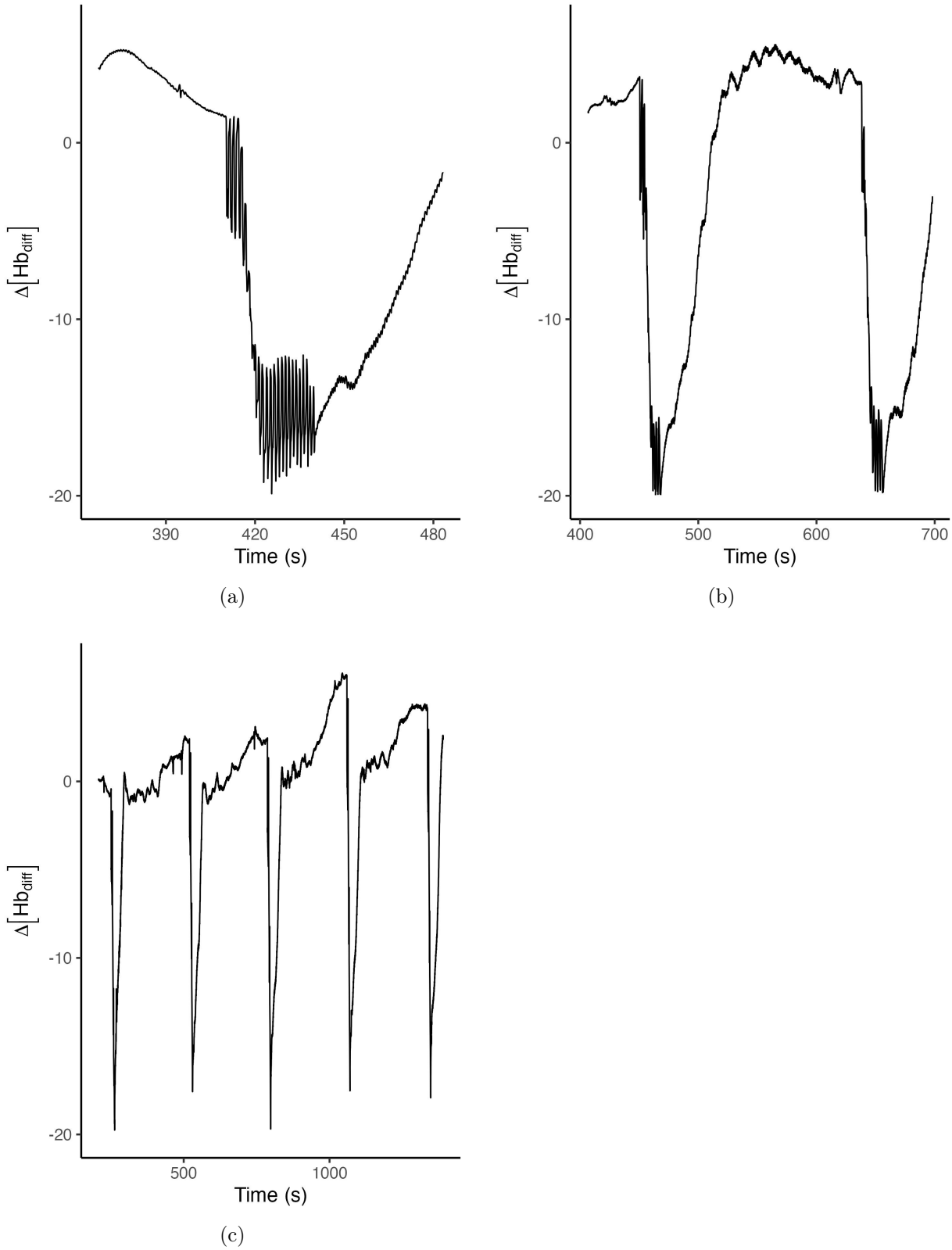


Figure 2.4. – Changes in $\Delta\,[\mathrm{Hb_{diff}}]$ over time during the three testing conditions $C1$, $C2$, and $C3$ for (a), (b), and (c), respectively.

Neuromuscular measurements

Activity of the three VLat, VMed and RFem was assessed through EMG using three sensors (Trigno Avanti, Delsys, MA, USA) located in respect of the SENIAM recommendations (Hermens, Freriks, Disselhorst-Klug, et al. 2000). Electrode sites were properly shaved and cleaned with alcohol before electrode placement. The sampling frequency of EMG signals was set at 2048 Hz, recorded through the software EMGworks (Delsys, MA, USA) and exported using the Delsys file utility application (Delsys, MA, USA).

Activity of quadriceps muscles were analysed in both time and frequency domains. In time-domain analysis, integrated signals amplitude were calculated from VLat, VMed and RFem for each knee extensions using a root mean square (RMS) function (see Equation 2.4), following a signal rectification and filtering using a second-order low-pass Butterworth filter with a cut-off frequency of 10 Hz. Then, normalisation to the mean signal computed from the first repetition and a time-normalisation were processed, ensuring unbiased analysis within testing sessions (Halaki and Ginn 2012).

$$f_{RMS} = \lim_{T\to\infty} \sqrt{\frac{1}{2T}\int_{-T}^{T} [f(t)]^2\, dt}\,. \tag{2.4}$$

In frequency domain analysis and since testing exercises involved dynamic contractions, STFT were processed on 125 ms overlapping samples of length $l = 250ms$. Then, power spectral density (PSD) representation (see Figure 2.5) enabled to extract MDF, as a valuable statistic for detecting impairment in EMG signals due to muscular fatigue (Ma'as, Azmi, et al. 2017; Phinyomark, Thongpanja, Hu, et al. 2012). MDF represents the frequency value at which the EMG power spectrum is divided into two regions with an equal integrated power. It is defined in the sequel :

$$\sum_{j=1}^{MDF} P_j = \sum_{j=MDF}^{M} P_j = \frac{1}{2}\sum_{j=1}^{M} P_j\,,$$

where P_j is the EMG power spectrum at a frequency bin j, and M is the length of frequency bin (Thongpanja, Phinyomark, Phukpattaranont, et al. 2013).

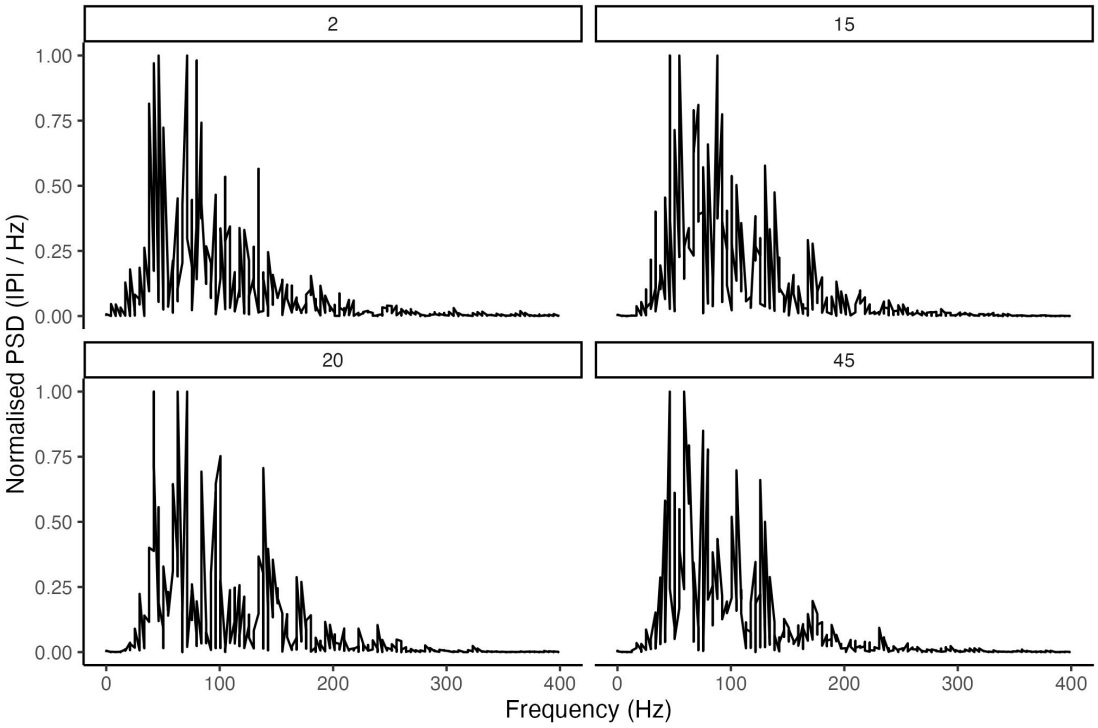


Figure 2.5. – Power spectral density of VLat calculated from STFT over four samples (from the second, fifteenth, thirteenth and forty-fifth samples of a single repetition).

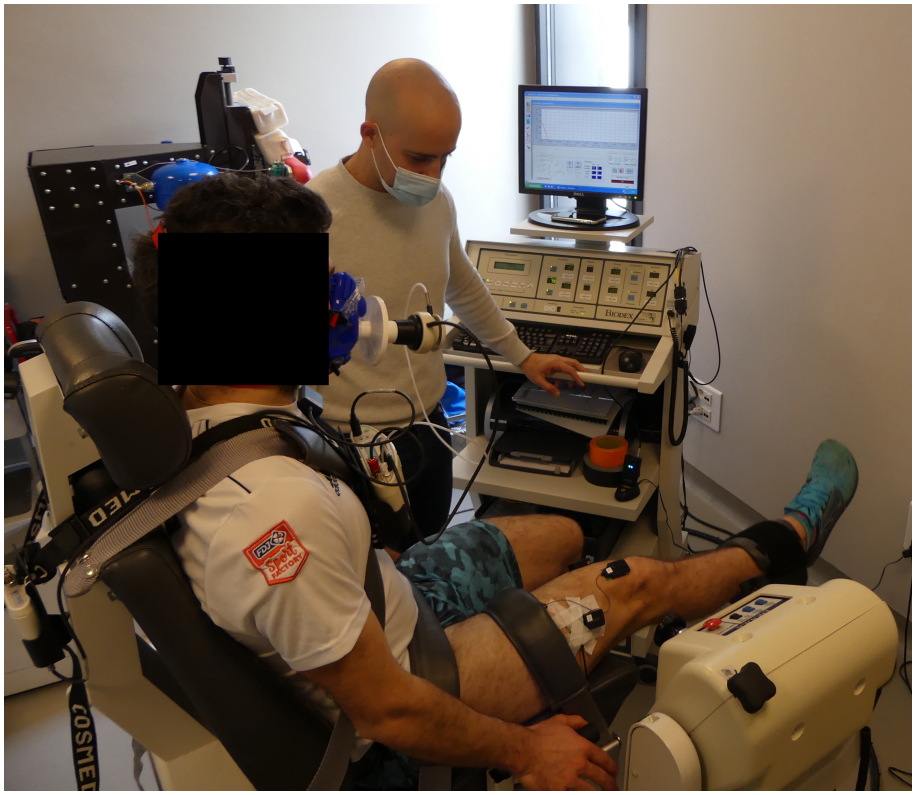


Figure 2.6. – Picture of a participant performing a testing session with all associated measurements.

2.1.4. Physiological responses to resistance exercises

In this section, we sequentially present the physiological responses to exercise, according to both resistance exercise sessions and individual T-V profiles.

Let us recall that participants performed three testing sessions at a relative theoretical intensity and RM, which differ in terms of number of set as well as inter-set recovery duration (see Sections 2.1.2.3 and 2.1.2.2 for details about how the testing sessions are volume-equated). Hence, we assume that in isokinetic conditions, each repetition is performed as a maximal exertion where the velocity of exercise gives intensity (i.e. the torque produced). On this basis, participants performed three, nine or twenty-four theoretical RM that should represent the last repetition before observing a significant decline of torque production. Furthermore, relying on individual isokinetic T-V profiles makes the relative intensity being directly related to TUT. Analysis were carried out through ANOVAs and hierarchical mixed models formulated in Appendix E.

2.1.4.1. Metabolic and hormonal responses to resistance exercises

Blood lactate concentrations.

A first assessment of metabolic responses to exercise shows that changes in $[lact_b]$ are mostly impacted by the testing session. As shown in Table 2.2, $C3$ was the testing condition that induced the lowest changes regarding baseline $[lact_b]$ ($\beta_2^* = -9.412 \in [-13.820, -4.853]\,95\,\%\,CI, p < 0.001\,, f^2(partial) = 2.66 \in [1.23, 4.58]\,95\,\%\,CI$ for mitigating effect when compared to $C1$). The moderate intensity protocol induced significant changes in $[lact_b]$, even being less than the low intensity protocol $C1$ ($\beta_1^* = -1.932 \in [-3.053, -0.803]\,95\,\%\,CI, p < 0.01\,, f^2(partial) = 0.10 \in [0, 0.40]\,95\,\%\,CI$ compared to $C1$). Also, there is a significant effect of the interaction between the testing conditions $C2$, $C3$ and $[lact_b]$ responses ($p < 0.05$). That means the effect of both $C2$ and $C3$ on $[lact_b]$ responses are not homogeneous across velocities, and higher is the velocity, lower is the change in $[lact_b]$ for a given testing protocol. These results are in line with literature since $[lact_b]$ are function of exercise intensity, but also of accumulation of exercise (i.e. repetitions) and inter-set recovery duration. Not surprisingly, sets of exercises performed at moderate to heavy intensity that account for a moderate to a large number of repetitions and short recovery duration (typically as performed in hypertrophy protocols) induced the greatest changes in $[lact_b]$ (Kraemer, Häkkinen, Newton, et al. 1999; Kraemer, Noble, Clark, et al. 1987; Marston, Peiffer, Newton, et al. 2017).

Table 2.2. – Parameters inference regarding changes in blood lactate concentrations in response to exercise. β^* represents standardised coefficient for each parameter of interest.

Effect	Parameter	β^*	Std.error	t	p.value	CI_{lower}	CI_{upper}
fixed	Intercept	3.273	0.519	6.306	< 0.001	2.287	4.240
fixed	C2	-1.932	0.599	-3.224	0.003	-3.053	-0.803
fixed	C3	-9.412	2.364	-3.982	< 0.001	-13.820	-4.853
fixed	velocity	-0.652	0.370	-1.762	0.086	-1.343	0.051
fixed	C2:velocity	-1.641	0.729	-2.252	0.031	-3.006	-0.282
fixed	C3:velocity	-7.814	2.400	-3.256	0.002	-12.303	-3.285
random	Intercept (sd)	0.762				0.257	1.251
random	Observation (sd)	1.025				0.752	1.255

Plasma cortisol concentrations.

Concentrations in plasma cortisol revealed that $C2$ and $C3$ lowered the $[cort_p]$ response to exercise, with a greater effect for $C3$ ($\beta_2 = -11.124 \in [-19.799, -2.449]\,95\,\%\,CI, p = 0.024\,, f^2(partial) = 3.75 \in [1.83, 6.29]\,95\,\%\,CI$ against $C1$). Exercise velocity also showed an lowering but negligible effect over the $[cort_p]$ response ($\beta_3 = -0.056 \in [-0.092, -0.020]\,95\,\%\,CI, p = 0.007\,, f^2(partial) \approx 0 \in [0, 0.03]\,95\,\%\,CI$) which looks homogeneous across testing session. Distributions of changes in $[cort_p]$ as well as parameter estimates of LMM are displayed in Figure 2.7 and Table 2.3.

Our results are supports ones reported by some authors who found that perform-

ing RE at heavier intensities induced the lowest change in $[cort_p]$ after exercise (Genner and Weston 2014). In another study, authors showed that an elevation of serum $[cort_p]$ was only observed after having performed moderate-intensity exercises with several repetitions and short inter-set recovery duration (70 % of the repetition maximum with a one-minute rest duration) (Kraemer, Dziados, Marchitelli, et al. 1993). It highlights the fact that changes in $[cort_p]$ depend on exercise intensity, but also exercise volume (number of repetition or duration of the set) and inter-set recovery duration. It can be seen in analogy with cycling exercises, where a short-term anaerobic exercise to exhaustion and prolonged aerobic exercises at high intensity has been shown to increase a cortisol response but not for exercises performed at very high intensities of short duration (Jacks, Sowash, Anning, et al. 2002; Kraemer, Patton, Knuttgen, et al. 1989).

In our case, $C2$ and $C3$ showed a decrease in $[cort_p]$ after exercise, but $[cort_p]$ remained unchanged in $C1$ compared to resting concentrations (see Figure 2.7). That indicates no testing sessions (neither $C1$, $C2$ nor $C3$) significantly induced an observable hormonal stressed state according to $[cort_p]$.

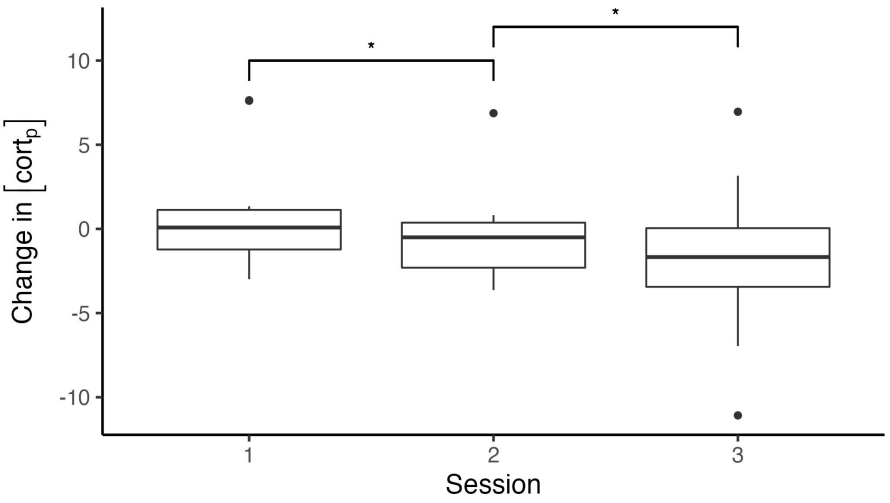


Figure 2.7. – Distributions of changes in serum $[cort_p]$ after having completed the three testing sessions. The asterisk denotes the level of significance $p < 0.05$ for differences between distributions.

In summary... These first results highlight that in lower-limb mono-articular exercise (knee extension), different lactate responses may occur for the same relative intensity. In this context, accounting for individual T-V or force-velocity profiles in the interpretation of acute metabolic responses is necessary. In addition, such exercises did not induce any significant hormonal stress according to the $[cort_p]$ responses to exercise.

Table 2.3. – Parameters inference regarding changes in plasma cortisol concentrations in response to exercise.

Effect	Parameter	β	Std.error	t	p.value	CI_{lower}	CI_{upper}
fixed	(Intercept)	9.383	3.714	2.526	0.016	2.519	16.246
fixed	C2	-9.410	4.926	-1.910	0.065	-18.514	-0.305
fixed	C3	-11.124	4.694	-2.370	0.024	-19.799	-2.449
fixed	velocity	-0.056	0.019	-2.904	0.007	-0.092	-0.020
fixed	C2:velocity	0.053	0.045	1.194	0.241	-0.029	0.136
fixed	C3:velocity	0.056	0.132	0.423	0.675	-0.189	0.301
random	Intercept (sd)	0.000				0.000	2.267
random	Observation (sd)	4.971				3.719	5.811

2.1.4.2. Cardiac kinetics and pulmonary gas exchange in response to resistance exercises

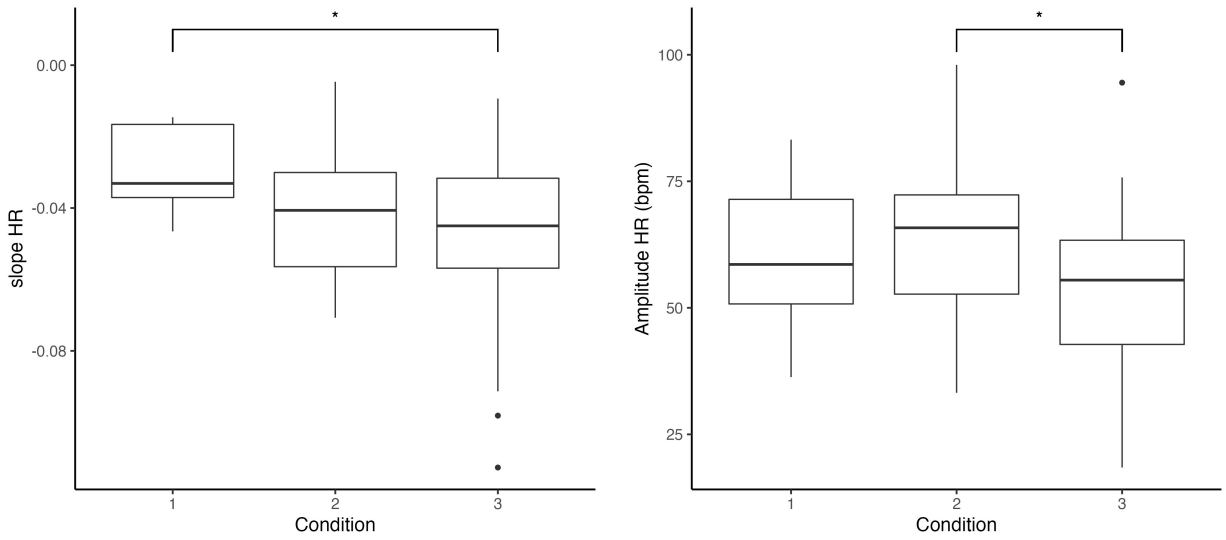
Heart-rate responses.

HR was continuously recorded during the testing sessions. Since RE are usually of short duration, we focused on HR amplitude ($A = max(x) - min(x)$) and recovery kinetics across exercises in order to estimate the cardiac activity demand.

As shown in Figure 2.8, one-way mixed repeated measures ANOVAs show that distributions of HR slopes significantly differ only between $C3$ and $C1$ ($\beta_{diff} = -0.018 \in [-0.033, -0.003]\, 95\,\%\, CI, p < 0.05\,, \eta^2 = 0.08 \in [0.01, 0.18]\, 95\,\%\, CI$). In addition, amplitudes of recovery computed from the onset of the recovery to the lowest resting value were significantly different only between $C3$ and $C2$ ($\beta_{diff} = -10.199 \in [-19.112, -1.286]\, 95\,\%\, CI, p < 0.05\,, \eta^2 = 0.09 \in [0.01, 0.18]\, 95\,\%\, CI$).

Changes in session-averaged HR slopes and accounting for individual variability (i.e. using random intercept parameter) indicate that both $C2$ and $C3$ were significantly related to the decrease of averaged HR slopes. However, the magnitude of testing condition effect over HR recovery kinetics remained very small ($f^2 \approx 0$). Regarding the changes in session-averaged amplitude during recovery, no relationship was found between the testing session and the outcome. Details are provided in Table 2.4.

These results indicate that globally, HR kinetics slightly differ between LI, MI and HI resistance exercises. When RE is performed at higher intensities, we observed shorter time courses of HR kinetics, be it appreciated in terms of session-averaged or for each recovery phase between series of exercises. However, differences in the magnitude of HR responses remain marginal. While $C1$ would increase HR at exercise over a large number of repetitions at LI, $C3$ results in similar changes through higher intensities and fewer repetitions.



(a) (b)

Figure 2.8. – Distribution of HR slopes and HR amplitudes across testing conditions. The asterisks denote the level of significance $p < 0.05$ for differences between distributions.

Table 2.4. – Parameter inference regarding changes in averaged HR slopes $\bar{\alpha}^{HR}$ and averaged amplitude of recovery $\bar{HR}_{amp}$.

Effect	Parameter	β	Std.error	t	p.value	CI_{lower}	CI_{upper}	DV
fixed	Intercept	-0.030	0.004	-6.847	< 0.001	-0.038	-0.021	$\bar{\alpha}^{HR}$
fixed	C2	-0.013	0.004	-3.229	0.004	-0.020	-0.005	
fixed	C3	-0.016	0.004	-3.835	0.001	-0.023	-0.008	
random	Intercept (sd)	0.013				0.008	0.020	
random	Observation (sd)	0.010				0.007	0.013	
fixed	(Intercept)	61.535	3.931	15.653	< 0.001	53.938	69.276	$\bar{HR}_{amp}$
fixed	as.factor(condition)2	1.930	4.216	0.458	0.651	-6.215	10.319	
fixed	as.factor(condition)3	-6.588	4.336	-1.519	0.142	-15.035	1.873	
random	Intercept (sd)	10.020				4.850	16.168	
random	Observation (sd)	10.442				7.698	13.698	

Oxygen uptake measurements.

As shown in Figure 2.3, $\dot{V}O_2$ was continuously recorded through the experiment.

At exercise, averaged rate of $\dot{V}O_2$ increase computed from linear relationships between $\dot{V}O_2$ and time indicated a slight trend for an improved rate of $\dot{V}O_2$ during

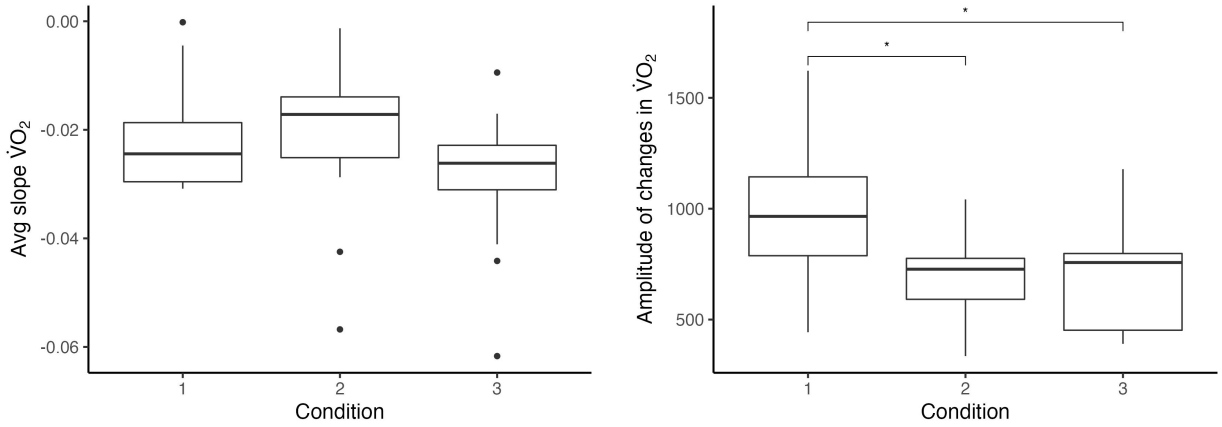
$C1$ only ($p > 0.05$). Nevertheless, this outcome has to be carefully balanced since $\dot{V}O_2$ measurements at exercises were greatly impacted by breath irregularities and apnoea times as encountered in some participants.

Post-exercise, session-averaged $\dot{V}O_2$ slopes computed from the mono-exponential function defined in Equation 2.3 show that distributions of exponential rate decays are not significantly different between $C1$, $C2$, and $C3$. On the other side, the session-averaged amplitudes for $\dot{V}O_2$ indicate greater $\dot{V}O_2$ amplitudes at recovery for $C1$ associated with higher maximal values of $\dot{V}O_2$ reached after exercise completion ($p = 0.017$, $\eta^2 = 0.23 \in [0.03, 0.41]\, 95\%\, CI$, see Figure 2.9). It suggests that performing less repetitions at higher intensities do not induce an elevation of $\dot{V}O_2$ at exercise and therefore no major changes in anaerobic metabolism contribution to the task completion, which supports precedent findings (Kang, Hoffman, Im, et al. 2005). It also corroborates the changes in $[lact_b]$ (being a witness of the anaerobic glycolysis contribution to the energy supply) for whose the changes were significantly higher for $C1$ then $C2$ than $C3$.

One may note that the $\dot{V}O_2$ value reached during the isokinetic –concentric– leg extensions tasks remain considerably lower than ones observed in poly-articular exercises and locomotor activities, mainly due to different muscle mass involved (MURAMATSU, KATAO, and HoMMA 1995). Also, the values reported in our experiment are consistent with the few studies that have investigated $\dot{V}O_2$ responses to resistance exercises (Farinatti and Neto 2011; Ratamess, Rosenberg, Kang, et al. 2014).

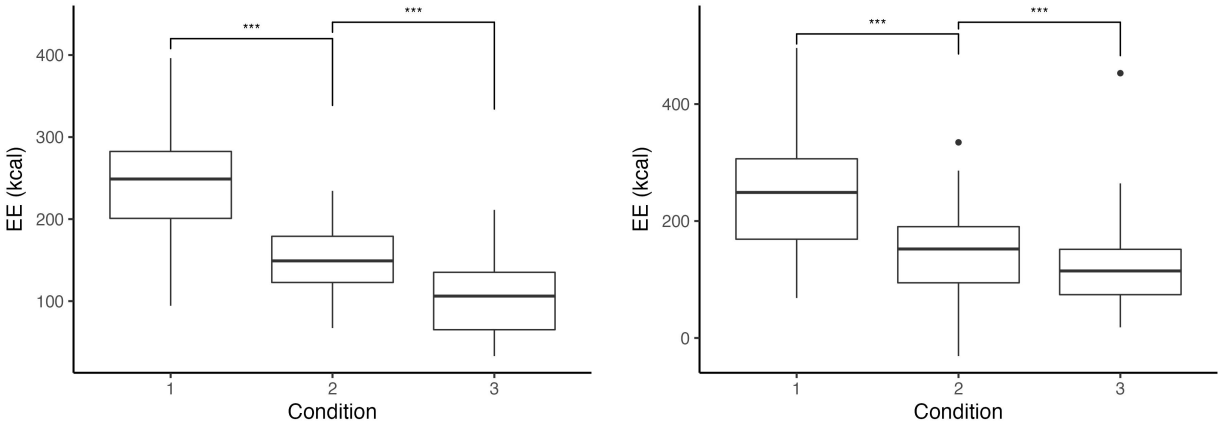
In addition, Figure 2.10 shows that even if testing sessions were balanced according to the VL's method, EE significantly decreases over the testing sessions. That was obviously expected since exercise velocity decreases drastically over testing sessions, therefore leading to an increase of the overall TUT. In addition, a greater amplitude of changes in $\dot{V}O_2$ is related to a greater total net EE, as shown in Figures 2.9b and 2.10a.

Also, differences in total EE (i.e. measured over exercise and recovery phases, Figure 2.10b) showed slightly lower differences between $C2$ and $C3$, compared to the EE measured on exercise phases only. This is in line with the expected negative correlation between inter-set recovery duration and acute $\dot{V}O_2$ responses, in which shorter recovery induces higher averaged $\dot{V}O_2$ and EE (Farinatti and Neto 2011; Ratamess, Rosenberg, Kang, et al. 2014).



(a) (b)

Figure 2.9. – Distribution of (a) session-averaged $\dot{V}O_2$ slopes computed over recovery phases and (b) session-averaged $\dot{V}O_2$ amplitudes. The asterisks * denote the level of significance $p < 0.05$ for differences between distributions.



(a) (b)

Figure 2.10. – Distribution of (a) energy expenditure from exercise phases only, and (b) energy expenditure from exercise and post-exercise revcovery phases. The asterisks *** denote the level of significance $p < 0.001$ for differences between distributions.

Since isokinetic leg extensions are mono-articular exercises and involve a few muscles, systemic measurements such as HR or $\dot{V}O_2$ might not correctly capture

the true metabolic demand –or grossly– of the exercise. Therefore, we completed these first analyses at a systemic level with analyses based on local measurements.

Muscular tissue oxygenation.

A first overview of the $\Delta\,[\mathrm{Hb_{diff}}]$ slopes computed at exercise and according to one-way repeated measures mixed ANOVA analysis, show that the distributions were significantly different between $C3$ and $C1$, and $C3$ and $C2$. Regarding the TSI slopes, only the means of distributions over $C3$ and $C2$ were significantly different (see Figure 2.11) but trends are in line with the results observed for $\Delta\,[\mathrm{Hb_{diff}}]$. In both cases, we found a significant and negative relationship between the physiological response and the exercise velocity ($\beta = -0.026 \in [-0.035, -0.017]\,95\,\%\,CI, p < 0.001$ and $\beta = -0.052 \in [-0.076, -0.030]\,95\,\%\,CI, p < 0.001$ for $\Delta\,[\mathrm{Hb_{diff}}]$ and TSI, respectively). However, not any significant differences were observed among distributions of $\Delta\,[\mathrm{Hb_{diff}}]$ and TSI amplitudes across testing conditions. ($p > 0.05$).

Given a set of rate constants α computed on the $\Delta\,[\mathrm{Hb_{diff}}]$ signal, we can identify $m\dot{V}O_2$ of the VLat for each series of N repetitions. In this case, $m\dot{V}O_2$ is determinate according to the slope α of the linear relationship between $\Delta\,[\mathrm{Hb_{diff}}]$ and time. As shown in Figure 2.11, $C1$ and $C2$ induced the greatest individual $m\dot{V}O_2$ responses, whereas $C3$ induced the lowest $m\dot{V}O_2$ responses. A reasonable explanation about lower $m\dot{V}O_2$ in $C3$ could be related to small ischemia-perfusion phenomena that appeared at lower velocities and within a single repetition. Hence, we differentiate the standard ischemia-perfusion observed at exercise - post exercise to jerky contractions highlighted in Figure (see Figure 2.12).

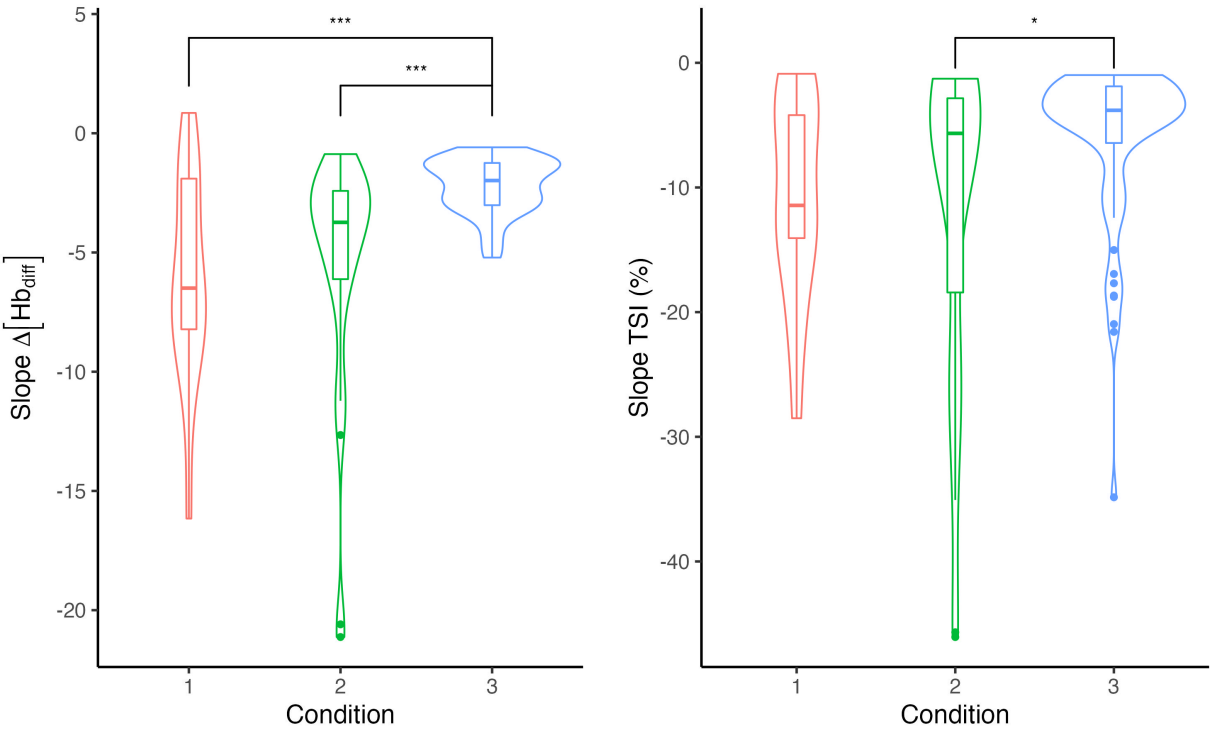


Figure 2.11. – Distribution of (a) $\Delta\,[\mathrm{Hb_{diff}}]$ slopes and (b) TSI slopes at exercise. Asterisks denote the level of significance for $*$, $**$, and $***$ being equivalent to $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

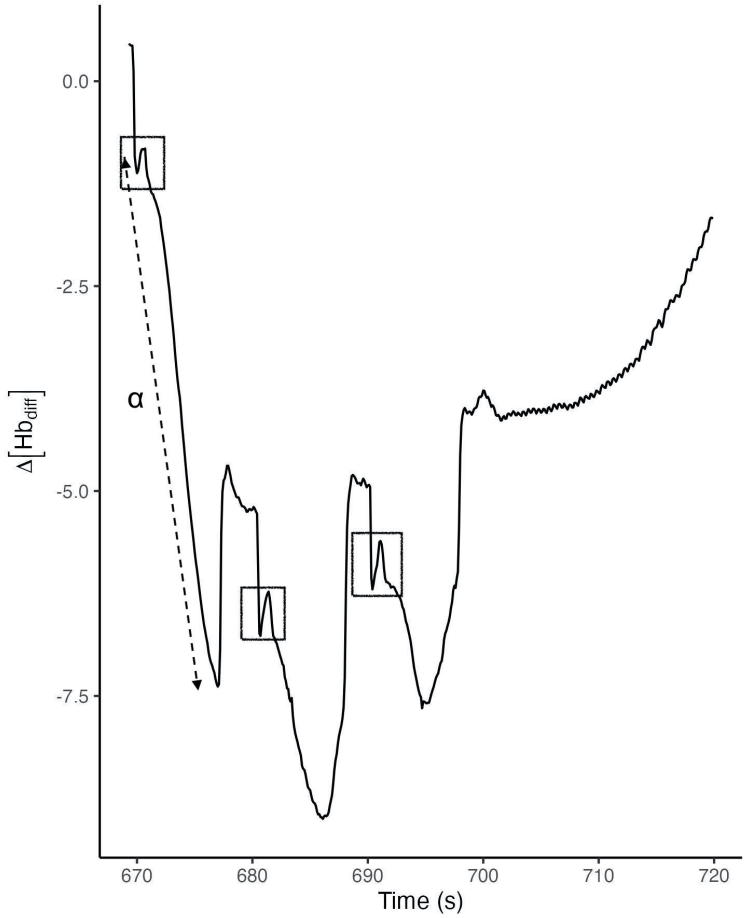
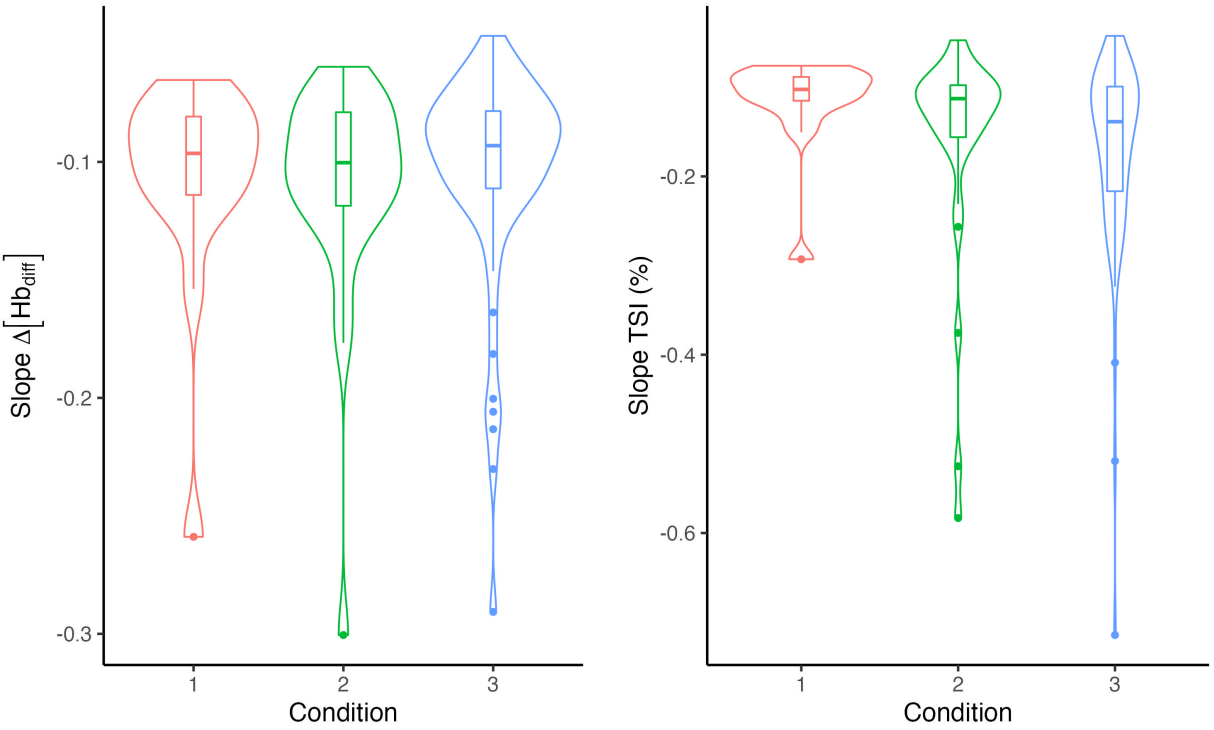


Figure 2.12. – Changes in $\Delta\,[\mathrm{Hb_{diff}}]$ over three isokinetic leg extensions performed at $0,174$ rad s^{-1}. For each repetition, short events of ischemia - perfusion phenomena are represented by squares. The slope of $\Delta\,[\mathrm{Hb_{diff}}]$ denoted α is computed over the longest –unbiased– interval of the first repetition.

Rate constants of $\Delta\,[\mathrm{Hb_{diff}}]$ and TSI computed over the half time of the recovery phases do not show any significant relationship neither with testing conditions, nor across exercise velocities ($p > 0.05$, see Figure 2.13). In addition, amplitudes calculated over $\Delta\,[\mathrm{Hb_{diff}}]$ and TSI during each recovery phases remain not significant different between testing conditions, and uncorrelated to the exercise velocity ($p > 0.05$). Hence, acute response of the vascular function seems to be invariant to the three levels of constraints imposed by $C1$, $C2$ and $C3$, respectively. An analysis of $\Delta\,[\mathrm{tHb}]$ confirmed that since there were no significant differences between $\Delta\,[\mathrm{tHb}]$ slopes over the three testing conditions ($p = 0.79$, see Figure 2.14). Therefore, recovery kinetics and amplitudes might be better suitable for monitoring chronic responses to exercise (*i.e.* over several sessions) rather than acute responses in healthy population and in such conditions.



(a) (b)

Figure 2.13. – Distribution of (a) Δ [Hb_{diff}] slopes and (b) TSI slopes during recovery phases.

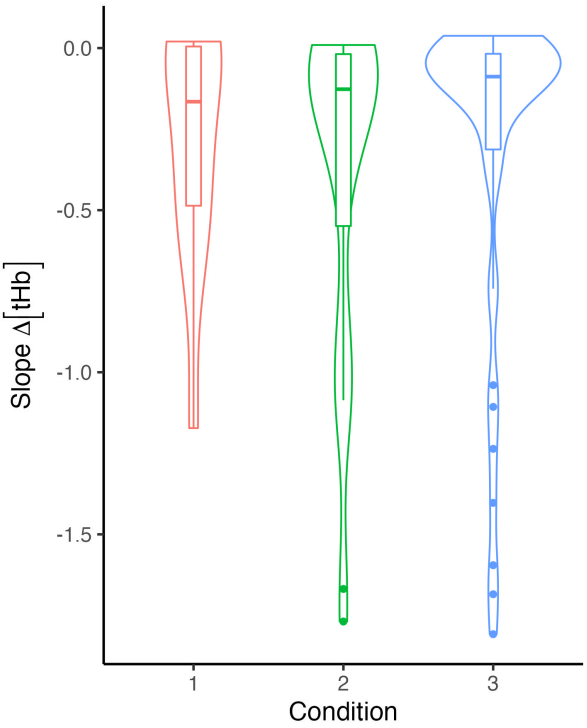


Figure 2.14. – Distribution of Δ [tHb] slopes during recovery phases.

2.1.4.3. Neuromuscular responses to resistance exercises

From mechanical measurements.

First of all, Figures 2.15b and 2.15a show that testing sessions imply significant differences in averaged normalised torque and total W_{mech}. As expected, $C3$ led to higher W_{mech} than the other sessions since lower exercise velocities imply higher TUT, and therefore a higher total W_{mech} for a same intensity. In addition, performing slower velocities allowed for higher torque values ($p < 0.001$). That is in line with the basic mechanisms of force production. Indeed, the contraction velocity is negatively correlated with the force production since the number of myosin heads activated and the amount of myosin heads in strong bound increases at slower contraction velocities, which result in higher production of force (Linari, Brunello, Reconditi, et al. 2015; Tyska, Dupuis, Guilford, et al. 1999).

An overview of averaged RFD and impulsion over testing sessions indicates that RFD_{peak} and RFD_{0-100} significantly increased between LI and MI ($p < 0.001$), but changes remain not significant between MI and HI sessions despite a large increase of exercise intensity (see Figure 2.15). These results are in line with the literature, suggesting that changes in RFD are partially related to exercise intensity (McGuigan and Winchester 2008; McGuigan, Winchester, and Erickson 2006). More generally, it has even been stated that RFD is not fully correlated with muscle strength (Guizelini, Aguiar, Denadai, et al. 2018) but also depends

on neuromuscular function properties and muscular determinants (e.g. muscle activation, motor units discharge rate, muscle fiber type composition and muscle architecture) (Maffiuletti, Aagaard, Blazevich, et al. 2016).

Meanwhile, impulsion showed significant changes across all sessions ($p < 0.001$). That was expected since impulsion is defined as the integral beneath the torque function and is highly dependant on TUT. Effects of the testing session on the measures presented in Figure 2.15 have been statistically tested using one-way repeated measure ANOVAs and Tukey Post-hoc tests.

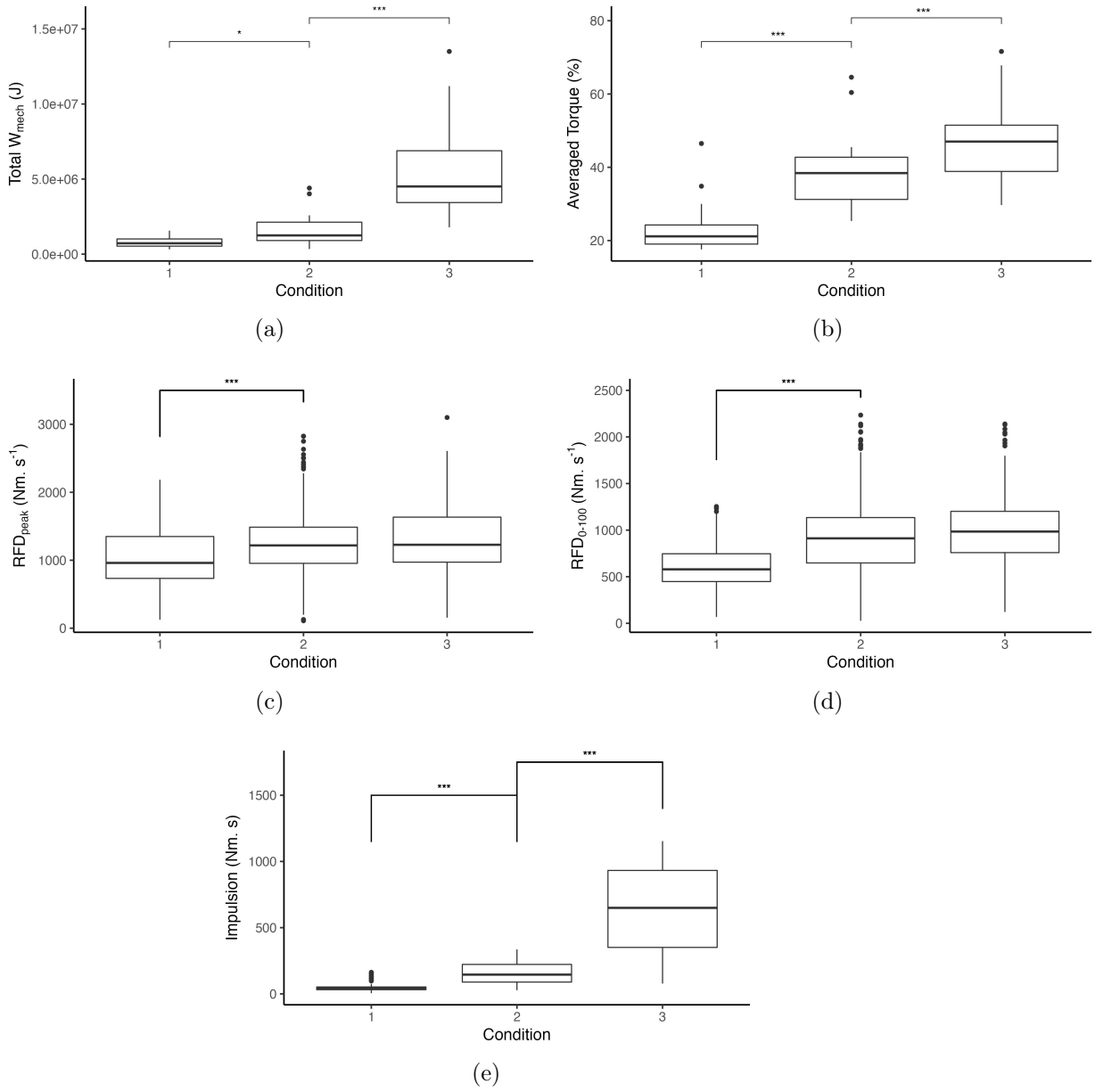


Figure 2.15. – Distribution of total W_{mech}, normalised averaged torque, RFD_{peak}, RFD_{0-100} and impulsion measures across testing sessions. Asterisks denote the significance level (*** being related to $p < 0.001$). In (c), (d) and (e), the distribution of the third session is significantly different from the first session (brackets are not displayed for clarity).

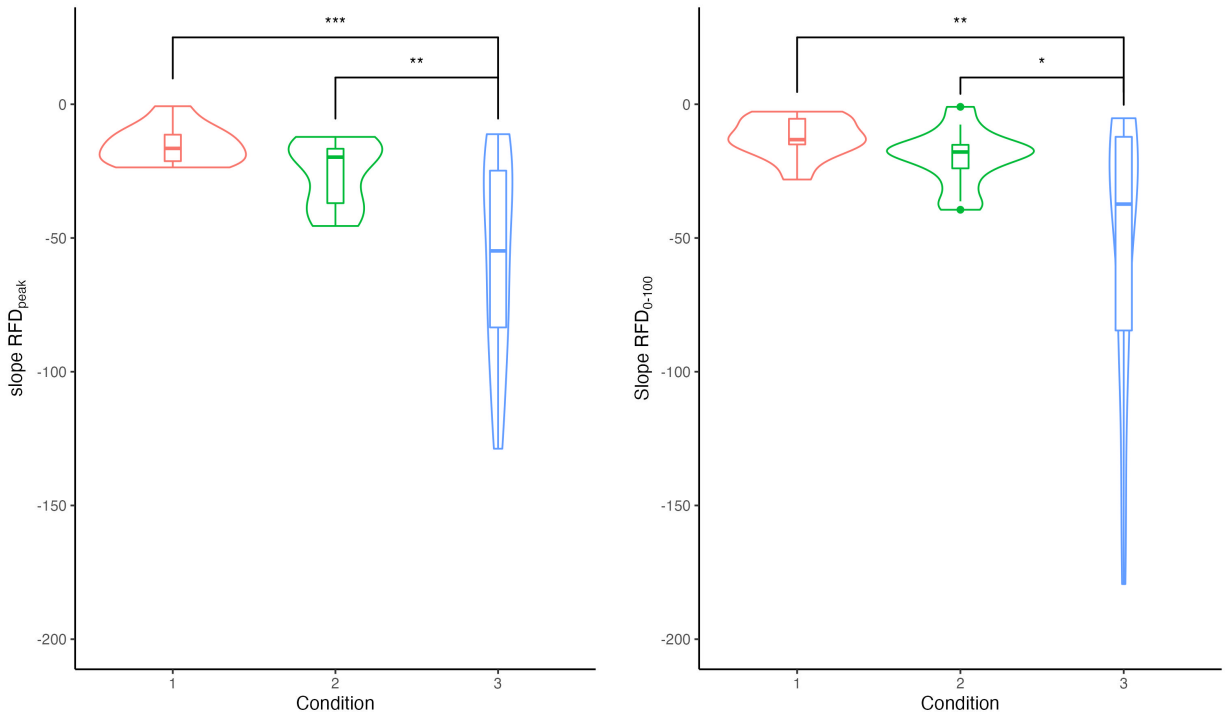
From this basis, it is possible to investigate neuromuscular impairments from torque measurements over each knee extension. As expected, performing numerous repetitions within or across series lowered both RFD_{peak} and RFD_{0-100}

($\beta_1 = -70.462 \in [-93.675, -47.242]\,95\,\%\,CI, p < 0.001\,, f^2(partial) = 0.16 \in [0.11, 0.22]\,95\,\%\,CI$ and $\beta_1 = -67.622 \in [-86.663, -47.775]\,95\,\%\,CI, p < 0.001\,,$ $f^2(partial) = 0.14 \in [0.09, 0.20]\,95\,\%\,CI$). However, the effect of repetition accumulation has to be considered regarding the testing condition since a significant interaction between the number of repetitions and the session suggest that slopes of relationships between RFD and the sessions differ. An explanation of that might directly comes from the protocol itself, since repetitions are allocated to a given number of sets (one, two and five sets for $C1$, $C2$ and $C3$, respectively), and where inter-set recovery duration differ between $C2$ and $C3$. Hence, frequency and duration of inter-set recovery influence changes in RFD and $C2$ shows a greater sustained RFD across repetitions than ones performed within $C3$. Details about model estimates are given in Table 2.5.

Having a decreasing effect of the repetitions accumulation over RFD_{peak} and RFD_{0-100} suggests that exercise induces progressive impairments of the neuromuscular function. Changes in impulsion did not evoke any decay through the time past at exercise and therefore, it does not represent the neuromuscular adaptations highlighted by both RFD_{peak} and RFD_{0-100}.

Since RFD linearly decreases over repetitions, it is possible to estimate the rate of fatigue apparition through individual regression slopes. That is in line with the work of Tesi, Colomo, Nencini, et al. 2000 where force generation (including rate constant of RFD) and inorganic phosphate release are closely related. The inorganic phosphate release corresponds to the shift in myosin heads and, therefore, a force production, for which changes in calcium concentration do not constitute a limiting factor of RFD (Allen and Westerblad 2001; Stehle, Solzin, Iorga, et al. 2009). Under muscle fatigue, ions H+ as well as inorganic phosphate concentrations increase in the myoplasm, therefore inhibiting the release of calcium in the sarcoplasmic reticulum and consequently the production of force (Allen and Westerblad 2001).

The low intensity condition $C1$ showed an homogeneous distribution of RFD and RFD_{0-100} rate decays. This suggests that the the apparition of fatigue was relatively consistent across participants. By contrast, a greater variability in RFD_{peak} and RFD_{0-100} rate decays across participants was found in $C2$ and $C3$, supporting the singularity in response to exercise at theoretical MI and HI. We note that according to the averaged population studied, distributions of RFD slopes are not significantly different between $C1$ and $C2$ ($p > 0.05$, see Figure 2.16).



(a) (b)

Figure 2.16. – Distribution of regression slopes for changes in (a) RFD_{peak} and (b) RFD_{0-100} across knee extension repetitions.

Table 2.5. – Parameters inference regarding RFD and impulsion responses to exercise.

Effect	Parameter	β^*	Std.error	t	p.value	CI_{lower}	CI_{upper}	DV
fixed	Intercept	1068.752	101.647	10.514	< 0.001	863.495	1274.027	RFD_{peak}
fixed	N_{rep}	-70.462	11.868	-5.937	< 0.001	-93.675	-47.242	
fixed	C2	204.758	20.886	9.803	< 0.001	163.902	245.622	
fixed	C3	248.123	23.733	10.455	< 0.001	201.698	294.554	
fixed	N_{rep}:C2	157.361	21.721	7.245	< 0.001	114.861	199.848	
fixed	N_{rep}:C3	113.672	26.616	4.271	< 0.001	61.598	165.734	
random	Intercept (sd)	389.970				272.613	566.575	
random	Observation (sd)	249.544				237.241	261.298	
fixed	Intercept	623.757	72.460	8.608	< 0.001	477.497	770.038	RFD_{0-100}
fixed	N_{rep}	-67.222	9.939	-6.763	< 0.001	-86.663	-47.775	
fixed	C2	321.079	17.492	18.355	< 0.001	286.862	355.303	
fixed	C3	392.394	19.876	19.742	< 0.001	353.512	431.279	
fixed	N_{rep}:C2	120.854	18.192	6.643	< 0.001	85.260	156.436	
fixed	N_{rep}:C3	73.603	22.291	3.302	0.001	29.989	117.203	
random	Intercept (sd)	276.984				193.345	402.723	
random	Observation (sd)	208.994				198.691	218.838	
fixed	Intercept	47.855	29.344	1.631	0.124	-11.282	106.989	impulsion
fixed	N_{rep}	-4.325	5.990	-0.722	0.470	-16.043	7.394	
fixed	C2	108.980	10.542	10.338	< 0.001	88.357	129.604	
fixed	C3	580.586	11.979	48.468	< 0.001	557.154	604.022	
fixed	N_{rep}:C2	1.153	10.964	0.105	0.916	-20.296	22.600	
fixed	N_{rep}:C3	15.585	13.434	1.160	0.246	-10.696	41.866	
random	Intercept (sd)	110.346				76.515	160.979	
random	Observation (sd)	125.955				119.745	131.888	

In summary, we found that :

— Any torque-related measurements were greater in the high-intensity protocol (*i.e.* $C3$) than the two others.

— Rate decays of RFD were also larger in $C3$. It suggests that high intensity RE induce a greater muscular fatigue according to RFD_{peak} and RFD_{0-100}.

Electromyographic activity.

Since EMG signals were normalised to the mean of exercises of interest, we investigated EMG activity or muscle activation patterns within participants and within sessions (see Figure 2.17).

In time domain, amplitude of EMG signal from knee extensors was given by the linear combination between RMS computed from averaged VLat, VMed and RFem signals. By assuming that resistance exercises might induce a muscular fatigue, we first investigated the effect of repeating sets (in $C2$ and $C3$) on a potential fatigue apparition. Using an intercept free LMM since $C1$ is filtered out due to a single set to be performed, we found that the averaged RMS computed across sets significantly decreased during $C2$ ($\beta_1 = -21.881 \in [-35.044, -8.558]\,95\,\%\,CI, p = 0.005\,, f^2(partial) = 19.803 \in [9.809, 33.152]\,95\,\%\,CI$). Also, the velocity of exercise showed a significant positive effect on changes in slopes of leg extensors RMS ($\beta_3 = 15.926 \in [4.125, 27.727]\,95\,\%\,CI, p = 0.018\,, f^2(partial) = 10.569 \in [5.051, 17.998]\,95\,\%\,CI$). The heaviest session, $C3$, did not show any significant change in RMS slopes. The results indicate that performing slower repetitions induces greater impairments of RMS, at least for exercises performed at MI. Non-significant changes reported in $C3$ might be supported by the fact that shorter sets coupled with longer inter-set recovery duration suggest a less effortful exercise. It allows for a better recovery after each set, despite intensities of exercise being higher than in other testing conditions (see Figure 2.15b).

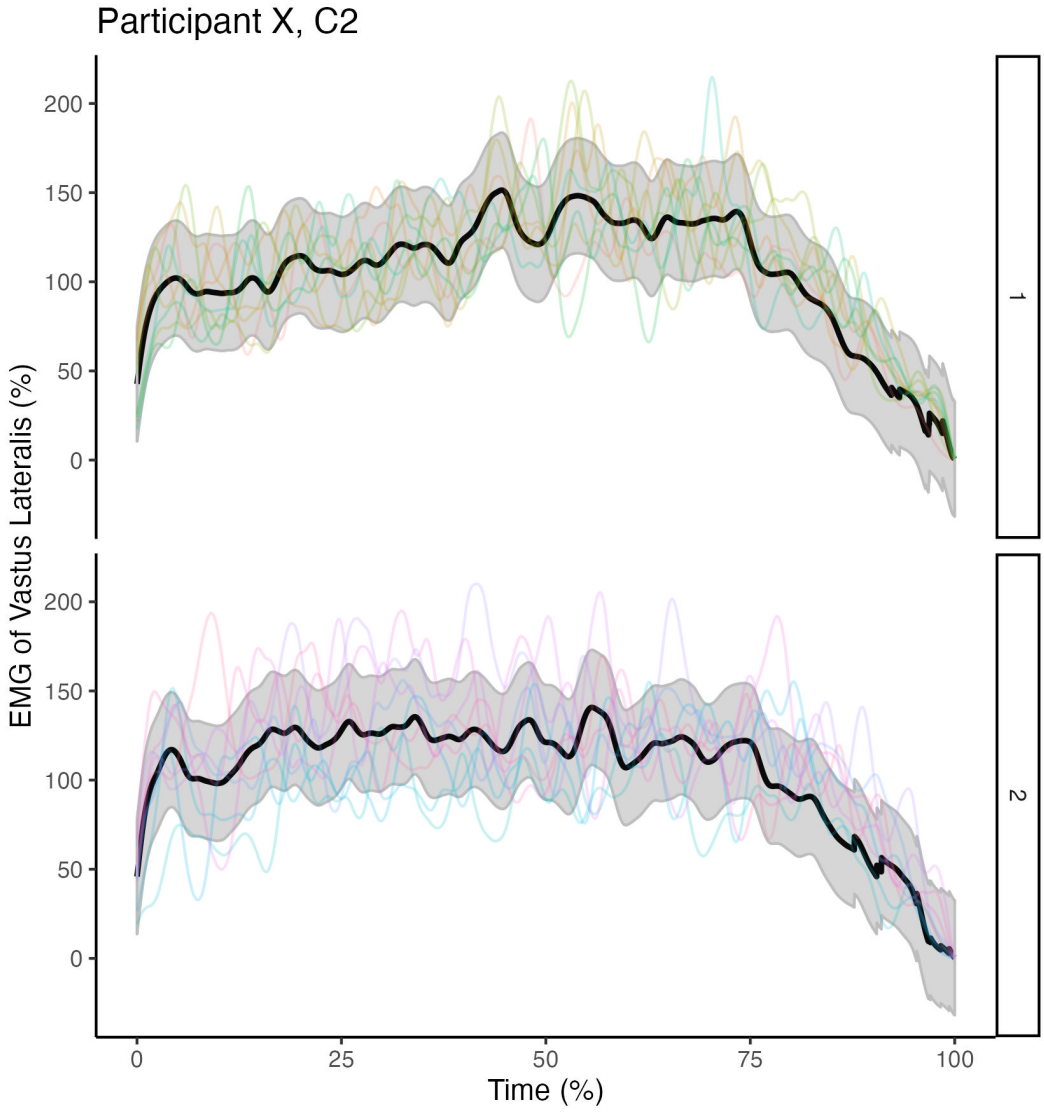


Figure 2.17. – Electromyographic signals of the VLat over two sets, recorded during $C2$ for a representative participant. The thin line represents the averaged signal surrounded by a one standard deviation ribbon and displayed over all repetitions.

On a closer inspection on changes in RMS over repetitions within sessions, we observed a small increase of RMS over repetitions ($\beta^* = 8.730 \in [4.690, 12.778]\,95\,\%\,CI$, $p < 0.001$, $f^2(partial) = 0.11 \in [0.06, 0.16]\,95\,\%\,CI$) and a slight decrease over normalised averaged torque ($\beta^* = -0.823 \in [-0.219, -0.428]\,95\,\%\,CI, p < 0.001$, $f^2(partial) \approx 0 \in [0, 0.01]\,95\,\%\,CI$). Both RMS and averaged normalised torque per repetitions are displayed in Figure 2.18.

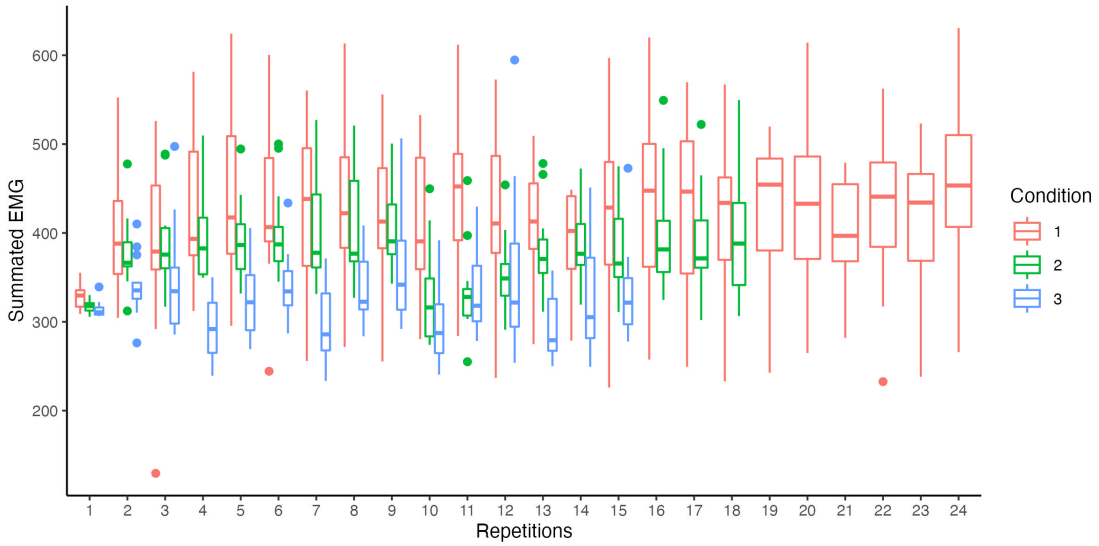
First, the decrease of normalised averaged torque was expected, in particular when the number of repetitions is large (Ebersole, O'Connor, and Wier 2006)

within a unique set (i.e.$C1$). By considering normalised averaged torque and RMS responses according to the session performed (i.e. the individual theoretical intensity), these results support the fact that the relationship between RMS and the torque produced is non-linear (or at least quadratic) (Kuriki, Mello, De Azevedo, et al. 2012; Madeleine, Bajaj, Søgaard, et al. 2001; Smith, Housh, Johnson, et al. 1998). An explanation of that might be related to (i) the fusion of individual motor units (MUs) and subsequent tetanus phenomena that occur between 60% and 80% MVC (Madeleine, Bajaj, Søgaard, et al. 2001) and (ii) the number and amplitude of recruited MUs are not directly related to changes in isokinetic exercise velocities (Smith, Housh, Johnson, et al. 1998).

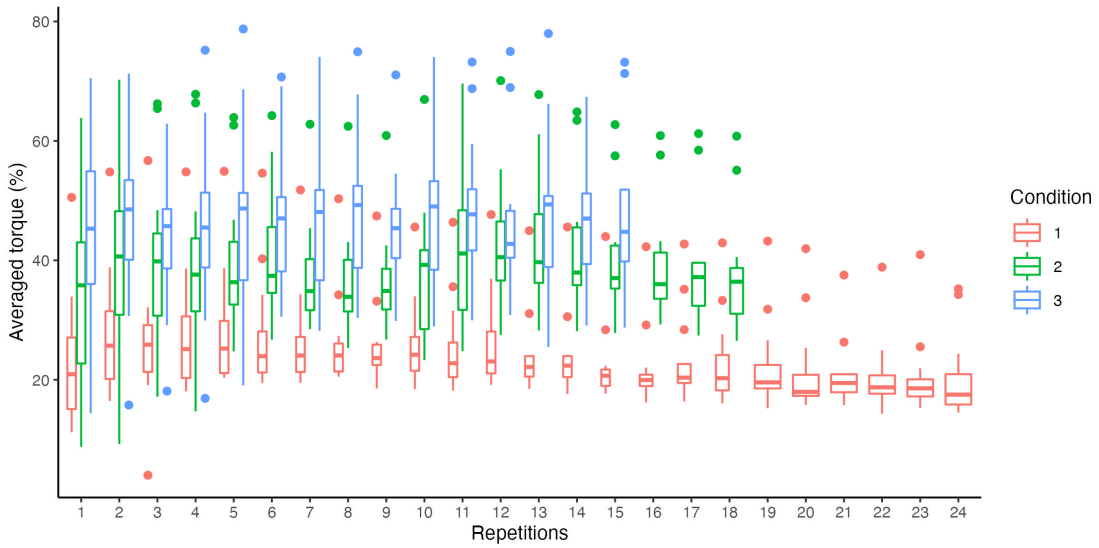
Even if the normalisation process of EMG signals that would not allow for proper inter-session and inter-participant comparisons (which is not the primary purpose here), the large and significant effect of exercise velocity over RMS responses is also in line with the literature (Smith, Housh, Johnson, et al. 1998). Also, a significant interaction effect between the testing session and individual velocities suggests that the relationship between RMS responses and exercise velocity is not homogeneous through testing sessions. Therefore, it can be considered –still in the light of the methodological limits– that EMG responses are likely better related to mechanical power rather than the torque produced, at least for exercise intensities up to moderate (Bodor 1999). Details about parameters inference of the discussed relationships are given in Table 2.6.

Table 2.6. – Parameters inference regarding the summated EMG signals (EMG_s) and the normalised averaged torque ($|\bar{T}|$) responses to exercise. Estimates β^* are standardised regression coefficients.

Effect	Parameter	β^*	Std.error	t	p.value	CI_{lower}	CI_{upper}	dependent variable (DV)
fixed	Intercept	376.969	13.272	28.402	< 0.001	351.232	404.090	EMG_s
fixed	N	8.730	2.068	4.221	< 0.001	4.690	12.778	
fixed	C2	4.528	13.171	0.344	0.732	-22.175	31.302	
fixed	C3	-50.893	59.541	-0.855	0.395	-172.570	70.940	
fixed	velocity	43.982	7.960	5.525	< 0.001	27.879	60.129	
fixed	C2:velocity	-51.639	11.390	-4.534	< 0.001	-74.226	-28.284	
fixed	C3:velocity	-48.122	47.279	-1.018	0.311	-143.725	48.820	
random	Intercept (sd)	38.178				24.999	56.029	
random	Observation (sd)	52.809				49.996	55.462	
fixed	(Intercept)	26.732	2.597	10.294	< 0.001	21.499	31.850	$\|T\|$
fixed	N	-0.823	0.202	-4.072	< 0.001	-1.219	-0.428	
fixed	C2	10.651	1.569	6.786	< 0.001	7.530	13.711	
fixed	C3	0.403	7.142	0.056	0.955	-13.791	14.328	
fixed	velocity	-1.907	0.947	-2.013	0.045	-3.791	-0.060	
fixed	C2:velocity	-6.504	1.265	-5.142	< 0.001	-9.010	-4.043	
fixed	C3:velocity	-17.897	5.526	-3.238	< 0.001	-28.869	-7.130	
random	Intercept (sd)	9.205				6.370	13.353	
random	Observation (sd)	5.160				4.884	5.418	



(a)



(b)

Figure 2.18. – Distributions of summated EMG signals from leg extensors and normalised averaged torque over repetitions.

Since leg extensions tasks are performed in isokinetic contraction (concentric only), frequency domain analysis might be more appropriate for revealing a potential fatiguing effect of exercises on the contractile function. Indeed, a downward shift of MDF_{quad} values over the three RFem, VLat and VMed muscles is observed through a significant decrease over the accumulation of repetitions ($\beta_3 = -2.778 \in [-3.520, -2.036]\,95\,\%\,CI, p < 0.001\,, f^2(partial) \approx 0 \in [0, 0.03]\,95\,\%\,CI$). In addition, significant decreases of MDF mostly occur during $C1$, but less in $C3$ then $C2$

(see Table 2.7 for details).

Table 2.7. – Parameters inference regarding MDF responses to exercise.

Effect	Parameter	β^*	Std.error	t	p.value	CI_{lower}	CI_{upper}
fixed	Intercept	69.077	2.012	34.329	0.000	65.020	73.134
fixed	C2	2.072	0.668	3.103	0.002	0.766	3.379
fixed	C3	1.099	0.759	1.448	0.148	-0.386	2.583
fixed	N	-2.778	0.380	-7.320	0.000	-3.520	-2.036
fixed	C2:N	3.027	0.695	4.357	0.000	1.668	4.385
fixed	C3:N	2.602	0.851	3.057	0.002	0.936	4.267
random	Intercept (sd)	7.600				5.280	11.078
random	Observation (sd)	7.980				7.587	8.356

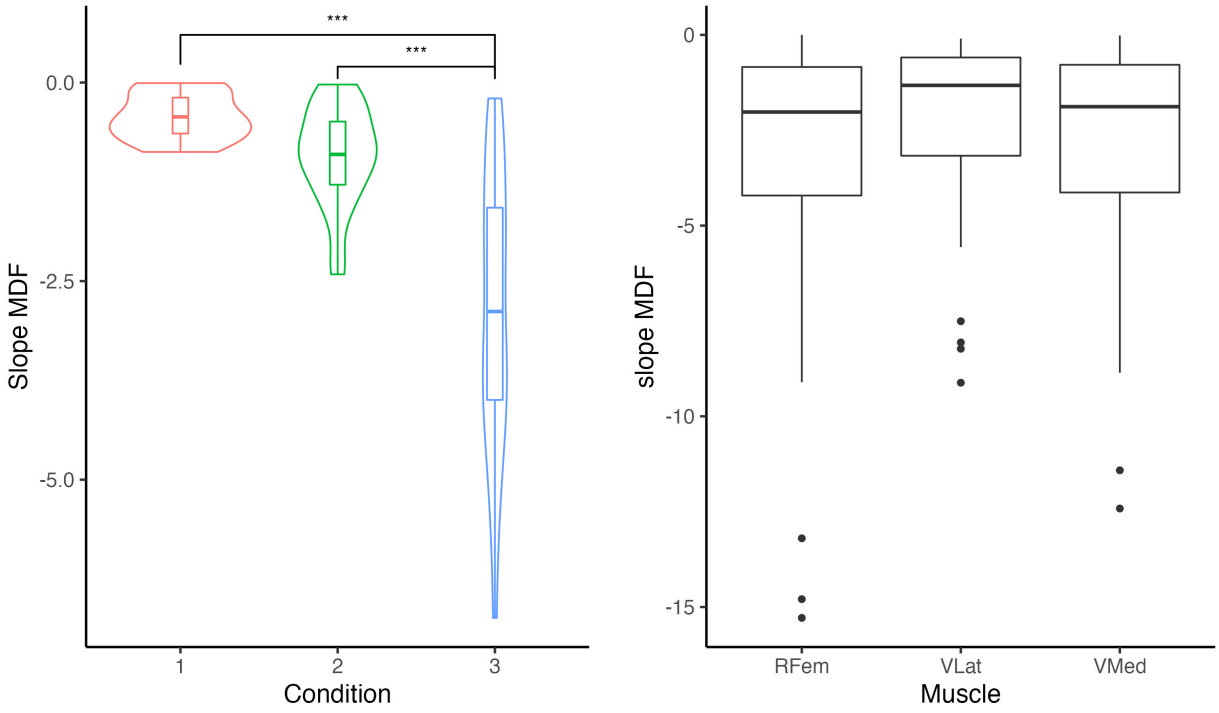
Then, MDF_{quad} computed from STFT samples let us compute slopes of MDF_{quad} decrease over repetitions within a set. Visually, distributions and dispersion of MDF_{quad} show a greater magnitude of muscle function impairments in $C3$ than $C2$ and $C1$ (see Figure 2.19a). The hierarchical LMM supports that finding with a greater and significant lowering effect of $C3$ over MDF_{quad} slopes than for the other testing sessions ($\beta = -2.197 \in [-3.106, -1.30]\, 95\,\% CI, p < 0.001\,, f^2(partial) = 0.06 \in [0.001, 0.22]\, 95\,\% CI$).

EMG activity analysis in the frequency domain brings insightful information about impairments of the contractile and nervous function, which could not be observed in the time domain. The results show that HI exercises, in our cases related to exercise velocity, induce the greatest magnitude of fatigue apparition and subsequent alteration of the muscle function. Specifically, mechanisms behind the decrease of MDF might be attributed to neurophysiological factors. First, a slower conduction velocity would be a reasonable explanation of MDF decrease over repetitions, according to the exercise properties (*i.e.* relative intensity, TUT).

In addition, neuromuscular responses we observed may be related to the muscle fibre type composition since it is accepted that phenotype muscle influences the force-velocity relationship (Nilsson, Tesch, and Thorstensson 1977; Thorstensson, Grimby, and Karlsson 1976; Thorstensson and Karlsson 1976). Indeed, the magnitude of MDF decrease may be correlated to the fibre type distribution, particularly the fast-twitch –type II– muscle fibres (Häkkinen and Komi 1983). Once again, RFem usually expresses a greater distribution of fast-twitch fibres and would be more prone to impairments by muscle fatigue than VLat and VMed. However, a muscle-by-muscle analysis revealed greater fatigue apparition on RFem, but not significantly different from VLat and VMed ($p = 0.22$). Distributions of MDF slopes are presented in Figure 2.19b.

Table 2.8. – Parameters inference regarding slopes of MDF_{quad} at exercise.

Effect	Parameter	β	Std.error	t	p.value	CI_{lower}	CI_{upper}
fixed	Intercept	-0.411	0.354	-1.161	0.249	-1.095	0.274
fixed	C2	-0.383	0.408	-0.939	0.351	-1.180	0.408
fixed	C3	-2.197	0.463	-4.742	< 0.001	-3.106	-1.300
fixed	Set:C	-0.022	0.038	-0.579	0.564	-0.096	0.052
random	Intercept (sd)	0.606				0.223	1.025
random	Observation (sd)	1.220				1.031	1.412



(a) (b)

Figure 2.19. – Distribution of regression slopes for changes in MDF across repetitions of isokinetic knee extensions, with (a) changes over conditions and (b) by muscle changes.

In Summary...

The neuromuscular analysis presented so far through mechanical measurements (i.e. RFD) and EMG (temporal and frequency analysis) shows that accumulating repetitions at LI, MI, and HI induced muscle function impairments. Through hierarchical LMM analysis, we observed within and between participants variability in terms of neuromuscular responses to exercise. That was expected according to

different modelled T-V profiles and, therefore, different exercise velocities and TUT across participants for the same relative exercise intensity (% MVC).

Globally, the accumulation of repetitions induced performance impairments which were associated with a fatigue apparition. Exercises performed at higher intensities showed a greater variability but also a greater magnitude of neuromuscular impairments. The results support that singularity in responses to resistance exercise predominates and should be considered in training prescriptions.

2.1.4.4. Conclusion

To conclude on the physiological responses to RE, we sum up the results as following:

— Concentrations in $[lact_b]$ were greatly impacted by the accumulation of repetitions within a series and the relative intensity, but also by the exercise velocity and the inter-set recovery duration. High exercise velocities for a given relative intensity induced lower changes in $[lact_b]$ concentrations.
— Heart rate kinetics slightly differed between testing conditions. Performing heavier exercise intensities induced shorter time courses of HR kinetics, also related to a greater amplitude of HR reached following exercise.
— Similarly, $\dot{V}O_2$ kinetics showed greater amplitude within $C1$ or LI testing sessions. As expected, metabolic EE was correlated with maximal values of $\dot{V}O_2$ reached at exercise.
— Locally, greater $m\dot{V}O_2$ were observed during $C1$ than $C2$ and $C3$. The TSI values supported that result, being correlated to $m\dot{V}O_2$.
— Rate of force development measurements (RFD_{0-100} and RFD_{peak}) indicated a greater fatiguing effect of HI testing sessions (i.e. $C3$).
— In the frequency domain, median frequencies confirmed the greater fatiguing effect of $C3$ compared to $C1$ and $C2$.
— The set of physiological adaptations to various RE associated to a sustained activity might reflect a muscle wisdom. As described by Enoka and Stuart 1992, the decline of torque and neuromuscular determinants across exercises may ensure an optimal production of force and an economical activation of fatiguing muscle by the central nervous system.

2.2. Towards a new model of training effect quantification

In Section 2.1.4, we presented physiological responses to three RE protocols that are initially volume-equated, according to the VL method and more generally, a RE based on percentage of RM. The intensity of exercises was drawn from individual T-V profiles in order to ensure that participants would perform RE according to their

own muscular properties. From the results presented so far, we found singularities in the physiological responses to RE, which were related to several parameters, including the *true* intensity (strain or stress) of exercise given by exercise velocity, as well as TUT and inter-set recovery duration. Hence, discretising the effects of training through objective TL should theoretically consider these parameters. In their study, Marston, Peiffer, Newton, et al. 2017 included inter-set recovery duration within some of the main TL quantification methods, such as based on RPE, VL or mechanical work calculations. It led authors to represent TL as a density of the effort done within a training session, which was better correlated to changes in $[lact_b]$ than the former methods. However, to our knowledge, no study has considered individual differences according to F-V profiles for quantifying such TL, yet being of importance in RT.

In a systemic approach, TL should be complemented by environmental, psychological, nutritional, technical and any factors that might influence an athlete status. Specifically, TL usually relies on an overall index that represents most of the adaptations induced by a training session (or a game, performance session). In that way, TL suffers from an empirical and restrictive dimension reduction, according to a few features –with more or less physiological meaning–aiming at discretising responses to exercise (Genner and Weston 2014; Marston, Peiffer, Newton, et al. 2017).

2.2.1. A physiological model of training load quantification in resistance training

Any TL quantification method should be as relevant as possible. In this way, Banister and Hamilton 1985 proposed a training quantification weighted by a non-linear weighting factor which hinges on the increase of $[lact_b]$ concentrations at various exercise intensities. However, their method mainly applies to endurance activities since the intensity was measured through HR. In RT, choosing an objective method of TL quantification mostly depends on empirical evidence. Even though the most common methods (see Section 1.1.1.2 for details) were partially correlated to some physiological responses to RT, they usually lack of physiological connotation in their structure.

In analogy to the bTRIMP from Banister and Hamilton 1985, we defined a new model based on neuromuscular impairments measured at different exercise intensities of resistance exercise. In this context, RFD appears to be (i) a relevant indicator of fatigue apparition and neuromuscular impairments according to the results presented in Section 2.1.4.3 and supported by the literature (D'Emanuele, Maffiuletti, Tarperi, et al. 2021; Häkkinen, Alen, Kraemer, et al. 2003), and (ii) a practical, raising and non-invasive parameter which benefits from the recent technological improvement of measurement systems (*e.g.* LPT and IMU). From averaged rate decays of RFD_{peak} observed at exercise, we can model the non-linear relationship between RFD_{peak} and exercise intensity (see Figure 2.20a), according

to a mono-exponential function (see Equation 2.3).

In addition, we consider a neuromuscular function being exponentially impaired by the accumulation of muscle fatigue. The ratio $\frac{1}{f(x)}$ describes that alteration such as represented in Figure 2.20b. Hence, we define a general formulation of a RFD_{peak}-based model of TL quantification in the sequel:

$$\begin{aligned} TL_{RFD} &= V\,I\left(\frac{1}{e^{\alpha\,I}}\right) \\ TL_{RFD} &= V\,Ie^{-\alpha\,I}\,. \end{aligned} \tag{2.5}$$

where V is the amount of repetitions performed, I denotes the relative intensity (% MVC) and α the rate decay such as $\alpha = -0.071$. From equation 2.5, we write its density correspondence

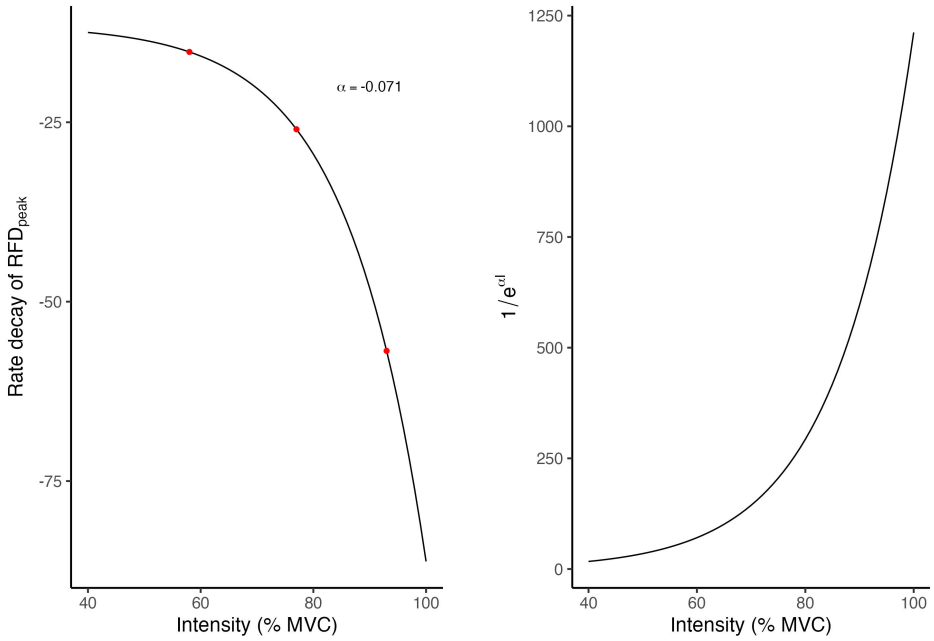
$$TL_{RFD_d} = \frac{TL_{RFD}}{R}\,, \tag{2.6}$$

with R being the total inter-set recovery duration.

Finally and like RFD, impulsion becomes a increasingly accessible measure nowadays. It could be used as a surrogate of the product of the volume and the number of repetitions. Hence, TL_{RFD} becomes

$$TL_{RFD^*} = \sum_{n=1}^{N}\int_{s=1}^{S} T(s)ds\; e^{-\alpha\,I}\,, \tag{2.7}$$

with N being the number of repetitions, S denotes the duration of each repetition and $T(s)$ is the torque produced.



(a) (b)

Figure 2.20. – Representation of (a) the non-linear relationship between the rate decay of RFD_{peak} and the relative intensity, and (b) the relationship between a neuromuscular breakdown given by the ratio $\frac{1}{f(x)}$ and the relative intensity.

Assuming that any TL quantification methods explain a part of the physiological adaptations, a generic equation of TL that accounts for more training related parameters or that combines several TL features might be a valuable alternative.

2.2.2. A linear combination of quantification methods and exercise related variables

Principal component analysis is a gold standard method for dimension reduction and multicollinearity issues. From a set of standardised explanatory variables, PCA projects them into a new space of orthogonal dimensions using a linear combination of the original variables, which are henceforth called principal components (PC). This makes PCA a valuable statistical method for dimension reduction problems, such as encountered when dealing with TL indexes.

From the results presented in Section 2.1.4, we processed a PCA based on training related features, the usual TL quantification methods and the three RFD_{peak} based models presented in Section 2.2.1. All these features are represented in Figure 2.21. The idea behind combining all these features into a PC is that PC would contain

information from each original feature, aiming at explaining a larger part of the total variance than isolated TL quantification methods.

The first two dimensions express 82.9% of the total data set inertia. That is, the first plane explains 82.9% of the individuals cloud total variance, which is significant since it is strongly higher than the reference value that equals 36.91 % (*i.e.* the 95% quantile of inertia percentage distribution of the first plane when the variables are fully independent). Note that among these dimensions, the first one explains the major part of the total variance (68.7 %, see Figure 2.22a).

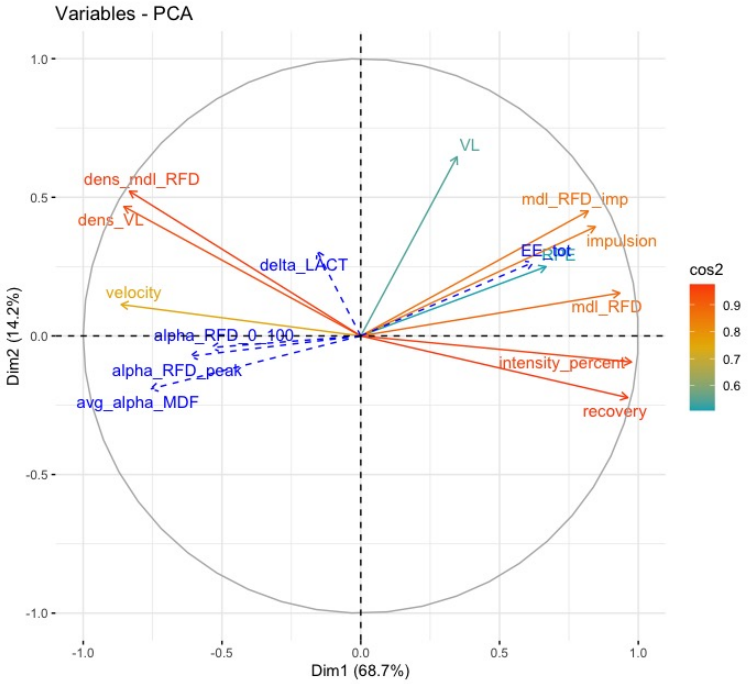
Graphically, the circle of correlation displayed in Figure 2.21a shows that most of the explanatory variables are represented alongside the first dimension. In contrast, VL seems to be represented on the second dimension. Among the main physiological responses highlighted in Section 2.1.4 retained for PCA analysis, RFD, MDF related ones as well as total EE are likely represented on the first dimension. Hence, a few key variables are identified for explaining these physiological responses such as ones that express the most representative within the first dimension (see Figure 2.22). Among these, TL_{RFD} (see Equation 2.5) seems to be a the most suitable TL quantification method. Changes in $[lact_b]$ seem to be barely explained and likely represented on the second dimension. At first glance, the density-like TL quantification methods (*e.g.* TL_{RFD_d} and VL weighted by the total recovery duration) appear to be the most explanatory methods of $[lact_b]$ changes. That makes sense since a longer recovery duration lower the accumulation of $[lact_b]$. Therefore, any TL quantification method that accounts for in the time past at recovery should better describe such metabolic response to exercise.

According to the testing condition, the contribution of each individual to the plane construction is identifiable (see Figure 2.21b). We retrieve some condition-specific physiological responses, especially for $C1$ and $C3$ for which the responses are generally diametrically opposed. The second testing condition usually sits in the middle of the $C1$-$C3$ interval, in which the physiological changes are moderate in regards to the other sessions. Concerning $C1$, the most characterising features are the density-related TL quantification methods and exercise velocity. In terms of physiological responses, we find the neuromuscular impairments indicators through MDF and RFD measurements. Once again, it supports the fact that the lower is the total recovery time, the higher is the impairment of the neuromuscular function.

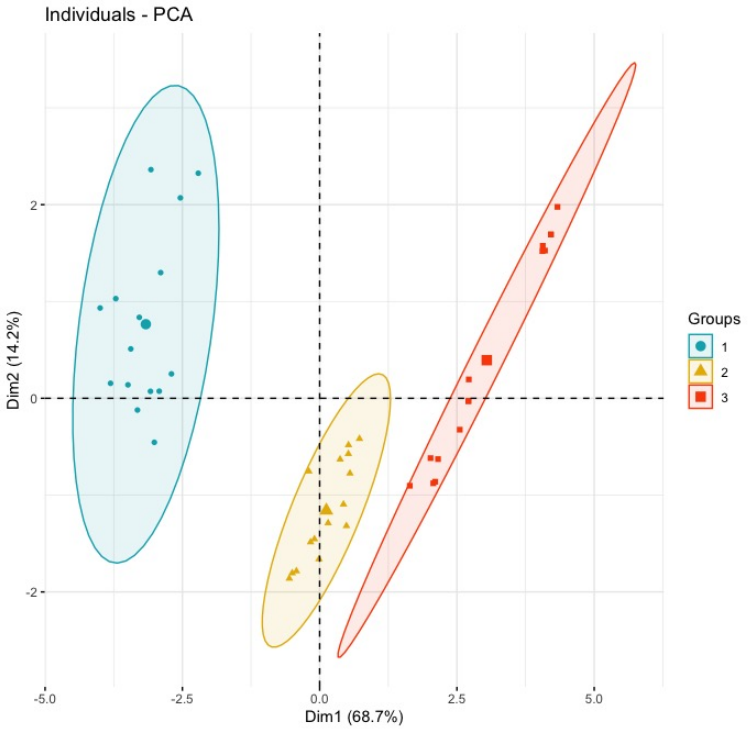
The second testing condition seems to be less explained by the first dimension. It is mainly represented by the recovery feature and is slightly related to changes in $[lact_b]$. Unlike the effect of recovery duration on neuromuscular changes, the recovery is negatively correlated with the increase in $[lact_b]$. Indeed, performing high-intensity RE and numerous repetitions associated with short recovery duration increases the contribution of the anaerobic lactic metabolism. Hence, we usually observe an increase of muscle capillarization and, therefore, an increased lactate clearance from the muscle to the blood (Kraemer, Deschenes, and Fleck 1988). By contrast, longer recovery duration inhibits such phenomenon and consequently lower the lactate response to exercise.

Regarding $C3$, it is likely represented on the first dimension. The main discriminating features are TL_{RFD}, TL_{RFD^*}, total impulsion, relative intensity, recovery and to a lesser extent RPE and VL. The positive correlation between these features and the total EE reflects the results presented in Section 2.1.4, where intensity and volume are the most determining factors of EE.

Finally, the recovery, relative intensity, $C1$ and $C3$ are highly correlated with this first dimension (respective correlation of 0.93, 0.95, 0.94, 0.97). These variables could therefore summarise themselves the first dimension, which could thus be used, through its coordinates, as an index of TL.

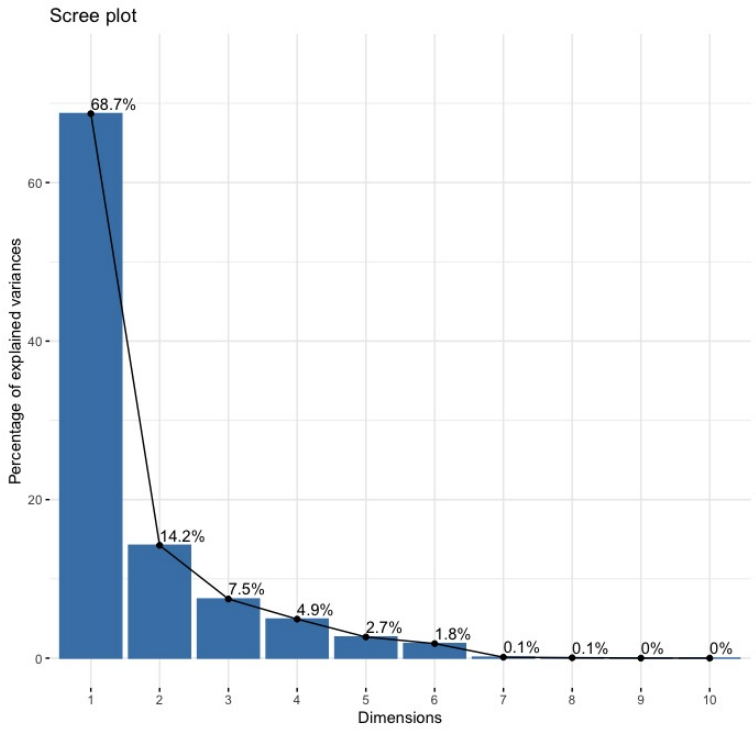


(a)

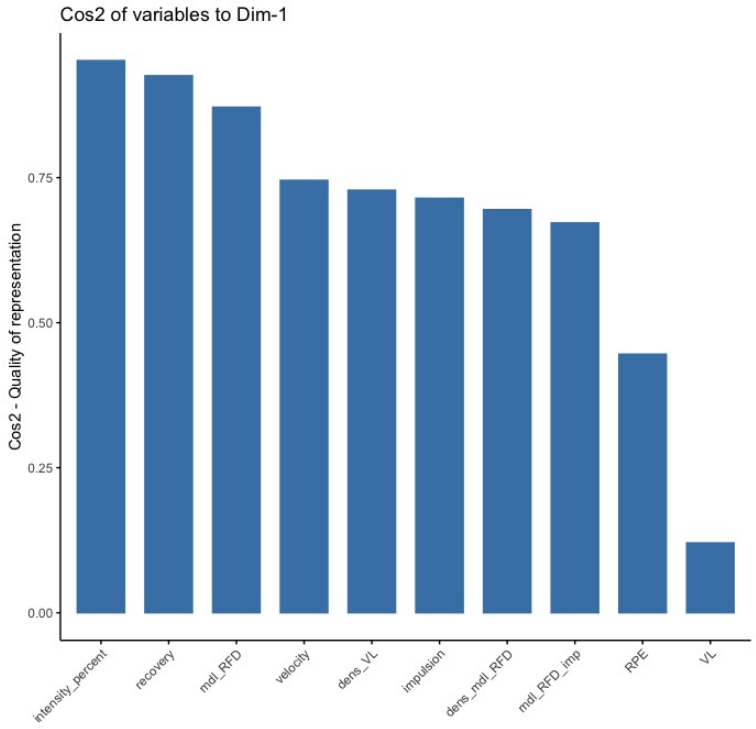


(b)

Figure 2.21. – Circle of correlations with (a) the quality of each variable representation (cos^2) and (b) the circle of individuals. In (a), variables in blue are illustrative variables, not accounted for in the calculation of distance between individuals. In (b), groups 1, 2 and 3 represents the testing sessions $C1$, $C2$ and $C3$, respectively.



(a)



(b)

Figure 2.22. – (a) Representation of principal components and (b) $\cos^2$ of variables to the first dimension.

2.2.3. Relationship between training load quantification methods and physiological responses

In order to validate the proposed models and to compare them with the usual TL quantification methods, we modelled the linear relationship between TL indexes and the main physiological responses observed at exercise using LMM with random intercepts. First, it appears that to weight TL estimate by the averaged RFD_{peak} kinetic provided greater correlation with any physiological responses to exercise. Indeed, the *physiological models* TL_{RFD}, its variant TL_{RFD^*} that uses impulsion or even the first PC provided better correlated TL indexes with physiological responses to exercise than ones quantified from the well-known RPE and VL methods (see Table 2.9).

At least two main reasons may explain these results. First, weighting a basic TL index by an exponential factor (according to the ratio presented in Equation 2.5) fitted over a heterogeneous population in various conditions make the TL being very sensitive to the exercise intensity, known to be a leading factor in responses to exercise. Then, PCA allows the compression of information from several features into PC (information drawn from each TL features and from training-related features, see Figure 2.22b). Therefore, it may better represent some physiological responses to exercise than if TL indexes were appreciated in isolation. Note that using PCA in such a way is not restricted to our set of measurements, but it can be processed in other contexts and data set for monitoring purposes.

Regarding the other TL quantification methods, the density representation of TL_{RFD} seemed to impair significantly the benefits provided by the *physiological* weighting term while smoothing the estimate across exercise intensities. However, our results are also in line with those from Marston, Peiffer, Newton, et al. 2017 insofar as the density representation of VL showed greater correlations than the simple VL.

In terms of subjective estimates of TL, RPE is outperformed by the objective proposed methods, but it remains better than the simple formulation of VL. This was to a certain extent expected since RPE is sensitive to RE intensity and velocity (Egan, Winchester, Foster, et al. 2006; Lagally, Robertson, Gallagher, et al. 2002).

Table 2.9. – Summary of the correlation analysis between four TL quantification methods and the main physiological responses.

Method	R^2	r	p.value	Target
RPE	0.49	0.70	0.98	$[lact_b]$
VL	0.52	0.72	0.39	
dens_VL	0.61	0.78	0.01	
Dim.1	**0.64**	**0.80**	0.01	
TL_{RFD}	0.60	0.78	0.01	
dens_TL_{RFD}	0.59	0.77	0.01	
TL_{RFD*}	0.55	0.74	0.23	
RPE	0.61	0.78	< 0.001	slope MDF
VL	0.07	0.26	0.08	
dens_VL	0.27	0.52	< 0.001	
Dim.1	0.67	0.82	< 0.001	
TL_{RFD}	**0.74**	**0.86**	< 0.001	
dens_TL_{RFD}	0.26	0.51	< 0.001	
TL_{RFD*}	0.71	0.85	< 0.001	
RPE	0.37	0.61	< 0.001	slope RFD_{peak}
VL	0.02	0.15	0.45	
dens_VL	0.37	0.60	< 0.001	
Dim.1	**0.58**	**0.76**	< 0.001	
TL_{RFD}	0.57	0.76	< 0.001	
dens_TL_{RFD}	0.35	0.59	< 0.001	
TL_{RFD*}	0.52	0.72	< 0.001	
RPE	0.29	0.54	< 0.001	slope RFD_{0-100}
VL	0.03	0.19	0.34	
dens_VL	0.27	0.52	0.01	
Dim.1	**0.42**	**0.65**	< 0.001	
TL_{RFD}	**0.42**	**0.65**	< 0.001	
dens_TL_{RFD}	0.27	0.52	0.01	
TL_{RFD*}	0.27	0.52	< 0.001	
RPE	0.54	0.74	< 0.001	EE
VL	0.03	0.19	0.29	
dens_VL	0.26	0.51	0.01	
Dim.1	0.68	0.83	< 0.001	
TL_{RFD}	0.72	0.85	< 0.001	
dens_TL_{RFD}	0.20	0.45	0.02	
TL_{RFD*}	**0.75**	**0.87**	< 0.001	

2.2.4. Conclusion

Investigating physiological responses to various RE showed singularities among participants. Indeed, elaborating RE protocols according to a percentage of RM should account for individual F-V profiles (T-V profiles in our case), considering the exercise velocity and by analogy, the TUT in order to understand the true demand of RE. That is also of significant importance for training monitoring purposes, where we aim at estimating the athlete response to training sessions based on a few features, namely TL.

The results found in the present study apply to local exercises in highly controlled conditions, in which participants performed concentric contractions of knee extensors in an open kinetic chain setting. Hence, the physiological responses highlighted in Section 2.1.4 might only represent a part of the overall responses underpinning RE performed in ecological conditions, with poly-articular and conventional exercises. Local responses (muscular) were accurately measured through gold standard measurement systems, but the systemic responses were probably underestimated.

In addition, heterogeneous T-V profiles were observed among participants (from exponential to likely linear profiles, see Figure 2.1). Aiming at representing the muscular properties of each participant and showing slight deviations from Hill's hyperbolic model, they may also differ from profiles modelled on poly-articular and multi-joint exercises, which remain mostly linear (Bobbert 2012; Jaric 2015; Rahmani, Samozino, Morin, et al. 2018). However, it does not discredit the use of F-V profiles for any training programming, quite the contrary, be it based on the RM or VBT method.

Ultimately, linear relationships between usual TL quantification methods and the overall responses were sequentially presented. We provided a new formulation of TL quantification based on physiological evidence. Using RFD as a physiological marker of the neuromuscular function, we extended the simplest formulation of VL to two exponential weighted models. Also, a third model was provided still according to a multiplicative exponential term but using impulsion as a surrogate of VL. Relationships with physiological responses to exercises showed that physiological-based formulations of TL indexes were better correlated than the usual objective and subjective methods.

Assuming that each of the usual TL index describes a part of some physiological responses to exercise, we showed that a dimension reduction method (e.g. PCA) that account for both usual, physiological-based TL indexes and other training-related features is a valuable alternative to the common TL indexes. In addition, the first component provided by PCA explains a large part of the total variance (*i.e.* nearly 70 %). It can thus be used as a surrogate for any unique index of TL in any monitoring processes for athletic performance or injury prevention purposes. Manifestly, the TL index based on PCA is not restricted to the set of variables we have provided in this study and could be extended to any other variables of interest.

The physiological-based TL features and their combination with other features of interest seem promising, but it remains to be investigated within conventional RT sessions in ecological conditions. In addition, compressing information into a single index of TL may be relevant in a practical sense. However, a multivariate approach should be encouraged for longitudinal analysis of athletic performance.

2.2.5. Take-home message

— T-V profiles have a strong implication in physiological responses to RE. It should be accounted for in any TL quantification, for example, through the measurement of exercise velocity.
— Conventional methods of TL quantification suffer from a lack of descriptive power according to physiological responses. New formulations of TL indexes based on rate decay of RFD showed a better relationship with the metabolic, neuromuscular and mechanical responses.
— Using PCA as a method for compressing information into a single feature benefits from the descriptive power of each TL quantification indexes as well as from other training-related parameters left aside. By accounting for the physiological-based formulation of TL, it could be used as a surrogate of the former TL indexes used so far.

3. Modelling responses to training loads and athletic performance prediction

Table of contents

3.1 Generalities . . . 138
3.1.1 Context . . . 138
3.1.2 Model selection and evaluation . . . 139
3.1.2.1 Holdout validation . . . 139
3.1.2.2 Leave-one-out cross-validation . . . 139
3.1.2.3 Leave-p-out cross-validation . . . 140
3.1.2.4 K-fold cross-validation . . . 140
3.1.2.5 Time-series cross-validation . . . 141
3.1.2.6 Model evaluation . . . 143
3.1.2.7 Take home message . . . 143
3.2 Training load responses modelling and model generalisation: application in short-track speed skating . . . 144
3.2.1 Data . . . 144
3.2.1.1 Population studied . . . 144
3.2.1.2 Definition of variables . . . 144
3.2.2 Modelling Methodology . . . 149
3.2.2.1 Reference model: Variable dose-response . . . 149
3.2.2.2 Regularisation procedures . . . 149
3.2.2.3 Random Forest . . . 151
3.2.2.4 Time series cross-validation and prediction . . . 151
3.2.3 Evaluation of model performances . . . 152
3.2.3.1 Models generalisation . . . 152
3.2.3.2 Accuracy of predictions . . . 155
3.2.3.3 Limits and conclusion . . . 158
3.3 Fitness-fatigue models: advantages, conceptual issues and contribution from machine-learning . . . 160
3.3.1 Making the most of control theory . . . 160
3.3.2 Current Fitness-fatigue framework and related issues . . . 161
3.3.3 A machine-learning perspective of the problem . . . 163

3.3.4 Take home message . . . 164
3.4 Using global positioning systems for modelling athletic performances 166
3.4.1 Methodological approach . . . 167
3.4.1.1 Population studied . . . 167
3.4.1.2 Data set definition . . . 167
3.4.2 Predicting the acceleration-velocity profile from GPS summarised features . . . 173
3.4.2.1 Time-series forecasting . . . 176
3.4.2.2 Multivariate modelling using commercial features . 179
3.4.2.3 Extracting new features from raw global positioning system data . . . 184
3.4.3 Conclusion . . . 187
3.4.4 Take home message . . . 188

3.1. Generalities

3.1.1. Context

In the previous chapters, we presented the most common methods for monitoring training loads and elucidating relationships between training and athletic performance. Two fundamental information arose, with (i) the use of objective and subjective measures of training loads and (ii) various mathematical modelling approaches, including statistical and ML methods. Despite the topic around athletic performance modelling being studied for many decades, there are still divergences in modelling approaches taken for somewhat the same purpose, such as optimising training programs for better athletic performance. Be the models linear with strong hypothesis (*e.g.* such as presented through FFMs models) or non-linear and hypothesis-free (*e.g.* various ML models), it is essential to being able to identify key factors of athletic performance within a variable environment. In addition, drawing physiological, psychological and practical interpretations of an athlete response to exercise from the models is highly expected. It would serve any coach or staff involved in an athlete accompaniment to the highest athletic performance level.

On this basis, interpreting models in terms of physiological patterns require having confidence in the models. Even models show great descriptive ability, they should foster the prediction of athletic performances under unknown situations. In other words, the ability of a model to generalise (i.e. see definition of generalisation (Bishop 2006)) becomes crucial. The generalisation is a statistical criterion which can be maximised through several procedures that are briefly presented –for the most common in ML– in the sequel (see Section 3.1.2). Ultimately, these procedures foster the determination of optimal models, relatively to the family of models considered and regarding their ability for generalisation. In the same time, they allow to diagnose model underfitting (Kouvaris, Clune, Kounios, et al. 2017) and overfitting (Lever, Krzywinski, and Altman 2016). While FFMs usually learned

parameters and make predictions from a single data set (with a very few exceptions) (Chalencon, Pichot, Roche, et al. 2015; Ludwig, Asteroth, Rasche, et al. 2019; Mitchell, Rattray, Fowlie, et al. 2020; Stephens Hemingway, Burgess, Elyan, et al. 2020), their generalisation ability cannot be ensured and further practical interpretation drawn from model parameters –such as time constants– might be flawed. In a predictive sense, nothing informs us about their accuracy in forecasting applications.

3.1.2. Model selection and evaluation

In this section, we review the most common methods used for model selection regarding generalisation problems.

3.1.2.1. Holdout validation

The simplest form of CV is the holdout method. It separates data set into two subsets, one for model training and the other for model validation, respectively. A target function is approximated from a function $f(x, \Theta)$ in a defined class of functions $\mathcal{H}$ (by definition, Θ denotes a set of hyper-parameters) that is fitted using the training data set. Then, output values are predicted on the validation data set. Splitting data in this way ensures that the data used for predictions have never been seen in the training step. An error criterion is reported (e.g. usually MAE or RMSE) for each prediction and averaged over the n observations, allowing for the model validation. This procedure lets us identifying an optimal function that minimises the criterion of interest, and therefore determines an appropriate set of model hyper-parameters for a given data set. Since data are separated into only two subsets, the holdout method has a low computation cost. However, the validation outcomes may suffer from a high variance because it depends greatly on data included in training and validation subsets (*i.e.* how the overall data set is divided).

3.1.2.2. Leave-one-out cross-validation

In order to produce a more robust model validation, many methods come with a more elaborated sub-setting of data. From a data set of n observations, LOOCV (Sammut and Webb 2010) implies to hold one observation and uses the rest of $n - 1$ observations to fit a model. Then, we evaluate the function on the withheld observation according to an error criterion. This procedure is performed n times until all the data are used for validation.

Since the model is trained on multiple subsets of data, LOOCV (and more generally CV) let to better estimate how the model will perform on unknown data. A drawback of LOOCV is related to a larger computation time than the holdout method. The LOOCV is presented in Figure 3.1.

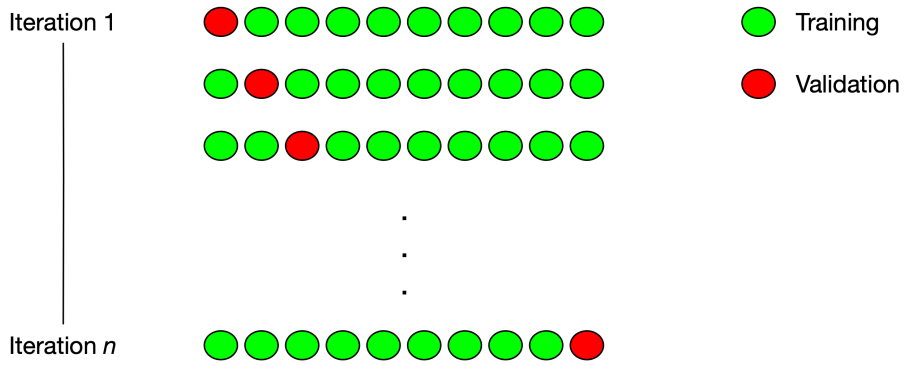


Figure 3.1. – Diagram of leave-one-out cross-validation. Green and red circles represent observations used for model training and validation, respectively.

3.1.2.3. Leave-p-out cross-validation

Like LOOCV, leave-p-out cross-validation (LpOCV) (see Figure 3.2) lets specify the number of observations p to be withheld for model testing. That way, the amount of testing iterations of a model can be determined using a mathematical combination $n \subset P$, leading to a more expensive computational task for large data set than LOOCV.

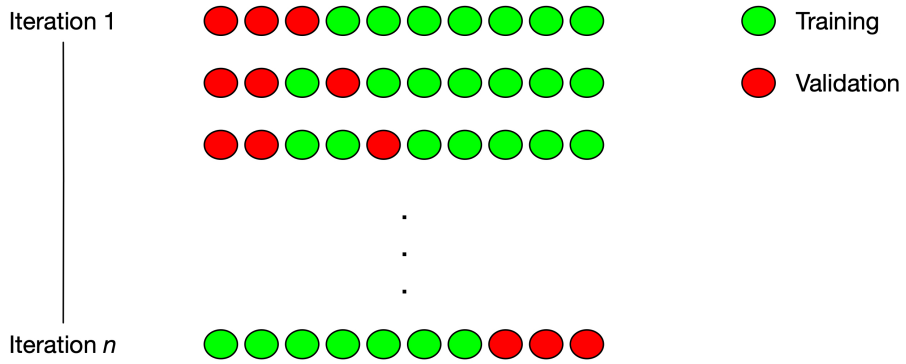


Figure 3.2. – Diagram of leave-p-out cross-validation, with $p = 3$. Green and red circles represent observations used for model training and validation, respectively.

Both LOOCV and LpOCV are exhaustive CV methods since models are trained and validated in all possible ways, using the full data set.

3.1.2.4. K-fold cross-validation

In a non-overlapping alternative, K-fold cross-validation (K-fold CV) (see Figure 3.3 comes by repeating the holdout method on k groups (i.e. folds) of the data

set. Folds come from a random division of the original data sample and remain approximately of equal size. The first fold is treated as a validation set, and the model is fitted on the remaining $k-1$ folds (James, Witten, Hastie, et al. 2013). k is fixed according to a value that lets samples be large enough to represent the broader data set statistically. Hence, increasing k leads to smaller subsets of the original data set, lowering the bias of the procedure. Note that there is a matter of a bias-variance trade-off behind the choice of k, according to the data set properties.

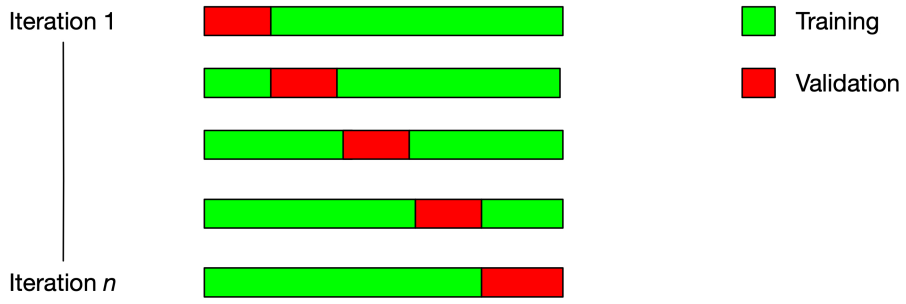


Figure 3.3. – Diagram of K-fold cross-validation, with $k = 5$. Green and red squares represent subsets of data (i.e. folds) used for model training and validation, respectively.

In comparison to K-fold CV, LOOCV and LpOCV remain particularly useful when working with small data set but incur performance trade-offs. LOOCV can also be seen as K-fold CV, where the number of folds is equal to the number of data points.

3.1.2.5. Time-series cross-validation

It is safe to say that athletic performance is related to time, including the past training sequences and the preceding athletic performances. It forms the basis of any FFMs presented in Section 1.2.1 in which the effects of training greatly depend on the training and performances that occur within a time window. From this point and in the case of frequently repeated athletic performances, there is a temporal dependency between athletic performances which means observations are not independent and identically distributed (i.i.d). Therefore, this temporal aspect has to be preserved all through the model selection and validation procedures. It also excludes any random re-sampling such as performed within K-fold CV, which might be no longer appropriate in this context.

Model selection procedures for time-series forecasting show a growing interest in statistics and ML (Bergmeir and Benı+tez 2012; Bergmeir, Costantini, and Benı+tez 2014; Bergmeir, Hyndman, and Koo 2018; Cerqueira, Torgo, and Mozetič 2020; Hyndman and Athanasopoulos 2018). In its simpler form and by analogy to the holdout procedure presented in Section 3.1.2.1, an out-of-sample (OOS)

procedure implies to re-sample training data into two time-ordered subsets, the first for the model training and the held out data (i.e. the second subset) for model validation.

Prequential alternatives to the simplest OOS allows for repeating model training and validation several times until all the training data are used. Resampling data according to the time involves blocks of sequential instances, wherein the initial iteration involves model training on the first block and validation using the second. In the next iteration, the second block is merged to the first block, thus increasing the size of the training subset, and the third block becomes the validation block. This is repeated until all the blocks are used for validation (see Figure 3.4a). The size of the validation block can be arbitrarily fixed according to the overall sample size and properties of the original training data set. This procedure is also sometimes known as an *evaluation on a rolling forecasting origin* since the *origin* at which the forecast begins rolls forward in time.

A variant of the aforementioned prequential procedure might be preferred, using a fixed but sliding window rather than a growing window (see Figure 3.4b). Performing either a growing or sliding window approach depends again on the structure and properties of the data set. For instance, a sliding window might be preferred to a growing window in the case of *large* sample size.

Finally, since there are temporal dependencies in the observations, performing any of the aforementioned approaches would not let fully independent observations be predicted. Strictly speaking, introducing a gap between the training blocks and the validation one would be welcomed and ensure a model selection and validation on fully independent data. That procedure –also called *hv - block cross-validation*– (Racine 2000) is presented in Figure 3.4c. The model selection would benefit from that multi-step errors, particularly if we are interested in forecasting tasks a few steps ahead. However, this alternative remains a little more data consuming that might raise small sample size issues in case of a limited number of available observations.

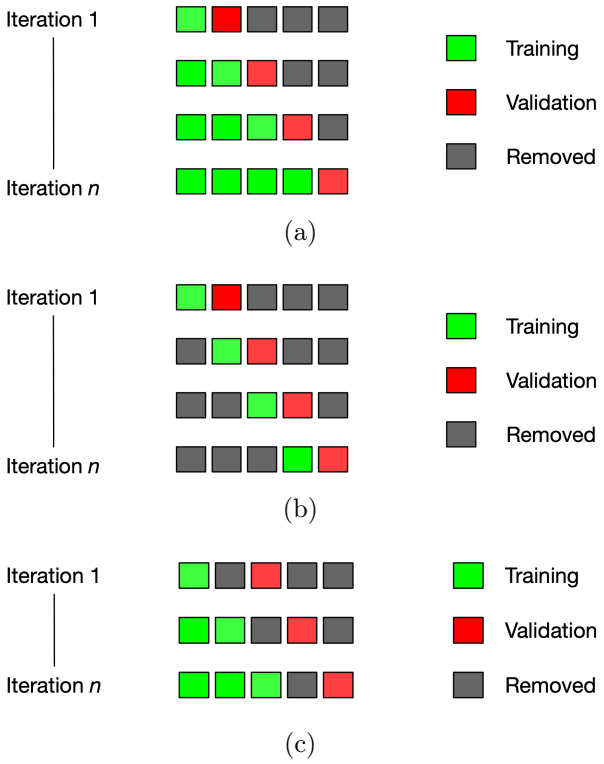


Figure 3.4. – Model selection according to prequential approaches. (a) represents a time-series block CV with a growing window, (b) represents a variant with a sliding window and (c) a variant with of a gap introduced within the sliding window approach.

3.1.2.6. Model evaluation

All along with the model selection procedure and according to the methods mentioned above, we seek an optimal model M^* that would provide the lowest errors on the validation subsets. That way, M^* is built to maximise performances on the validation data, but not necessarily on unknown and unbiased data. In order to address such validation bias issue, we after evaluate M^* on held-out testing data that were kept fully hidden from any step of the model selection procedure.

3.1.2.7. Take home message

— Several methods foster the determination of optimal models regarding their generalisation capabilities. Their choice relies on the structure of the available data, such as the sample size and time dependencies.
— A robust evaluation of optimal models lies on a training data set involved in the model selection process (*i.e.* training and validation data) and a testing data set made up of unknown data for evaluating the model performances.

3.2. Training load responses modelling and model generalisation: application in short-track speed skating

Physiological adaptations to exercise are complex. A multivariate approach of the problem would be appropriate to catch meaningful information from a set of features and, hence, to elucidate the relationship between training and athletic performance. However, it remains unclear whether a particular model family (i.e. physiological-based FFMs, statistical and ML models) should be preferred for athletic performance predictions and training protocol optimisation. This is supported by a lack of evidence and confidence in the modelling of training effects over athletic performance and the generalisation capabilities and accuracy of models.

In the sequel, we provide a modelling methodology that relies on model generalisation in the context of athletic performance modelling. Using data of elite athletes, we considered the DR (Busso 2003) as a baseline framework and compared it to three models: a PCR, an ENET regression and a random forest (RF) regression model. These models allow us to present and discuss the help of regularisation methods regarding the generalisation concept.

3.2.1. Data

3.2.1.1. Population studied

Seven national elite Short-track speed skaters (mean age 22.7 ± 3.4 years old; 3 males, body mass of 71.4 ± 9.4 kg, and four females, body mass of 55.9 ± 3.9 kg) voluntary participated in the study. Each athlete experienced the 2018 Olympic Winter Games in PyeongChang, South Korea ($n = 2$) or was preparing for the Olympics Games of Pekin, China ($n = 7$). The whole team was trained by the same coach, responsible for training programming and data collection. The mean weekly volume of training was 16.6 ± 2.5 hours. Data were collected over a three-month training period without any competition, interrupted by a two-week break and beginning one month after resuming training for a new season.

3.2.1.2. Definition of variables

Dependent variable: athletic performance. Participants performed each week standing start time trials (distance = 166.68 meters equal to 1.5 lap) after a standardised warm-up. At the finish line, the timing gates system (Brower timing system, USA) recorded individual time trial performance in order to ensure a high standard of validity and reliability between measures (Bond, Willaert, and Noonan 2017; Bond, Willaert, Rudningen, et al. 2017). A total of $n = 248$ performances were recorded for an average of 35.4 ± 2.23 individual performances. The performance

test being a gold standard for the assessment of acceleration ability (Felser, Behrens, Fischer, et al. 2016), athletes were all familiar with it before the study.
In the sequel, let $\mathcal{Y} \subset \mathbb{R}$ be the domain of definition of such a performance and $Y \in \mathcal{Y}$ the random variable. In this context, each observation $y_j \in Y$ can be referenced by both its athlete i and its day of realisation t as $y_{i,t}$.

Independent variables. Independent variables stem from various sources, which are summarised in Table 3.1. In the sequel, let $\mathcal{X} \subset \mathbb{R}^d$ with $d \in \mathbb{N}$ be the domain of definition of the random variable $X = [X_1, \ldots, X_d] \in \mathcal{X}$. The variable X is thus defined as a vector composed of the independent variables detailed hereafter. First, $\{X_1\}$ refers to the raw TL (see Figure 3.5c), calculated from on-ice and off-ice training sessions (see Appendix F). Then, $\{X_2, X_3\}$ represent two aggregations of daily TL. Those aggregations come from the daily training loads $w(t)$ –also known here as X_1– convoluted to two transfer functions adapted from Philippe, Borrani, Sanchez, et al. 2019, which are denoted $g_{\text{imp}}(t)$ and $g_{\text{ser}}(t)$.

The associated impulse response $G_{\text{imp}}(t)$ reflects the acute response to exercise (e.g. fatigue), according to Equation 1.9. It is defined as

$$G_{\text{imp}}(t) = e^{\frac{-t}{\tau_I}} \,, \tag{3.1}$$

where τ_I is a short time constant equals to 3 days in this study (Figure 3.5a). Respectively, the response $G_{\text{ser}}(t)$ describes a serial and bi-exponential function reflecting training adaptations over time. It is defined as

$$G_{\text{ser}}(t) = \left(1 - e^{\frac{-t}{\tau_G}}\right) U \; + \; e^{\frac{-(t-TD)}{\tau_D}} \; | \, U - 1 \, | \,, \quad \text{with} \quad U = \begin{cases} 1 & \text{if} \quad t < TD \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

The time delay for the decay phase to begin only after the growth phase is given by the constant TD. Here, $TD = 4\tau_G$. Both τ_G and τ_D are the time constants of respectively the growth phase and the decline phase with $\tau_G = 1$day and $\tau_D = 7$days (Figure 3.5b). Note that the time constants τ_I, τ_G, τ_D were averaged and based on empirical knowledge and previous findings (Busso 2003). Hence, for a given athlete and according to Equation 1.13,

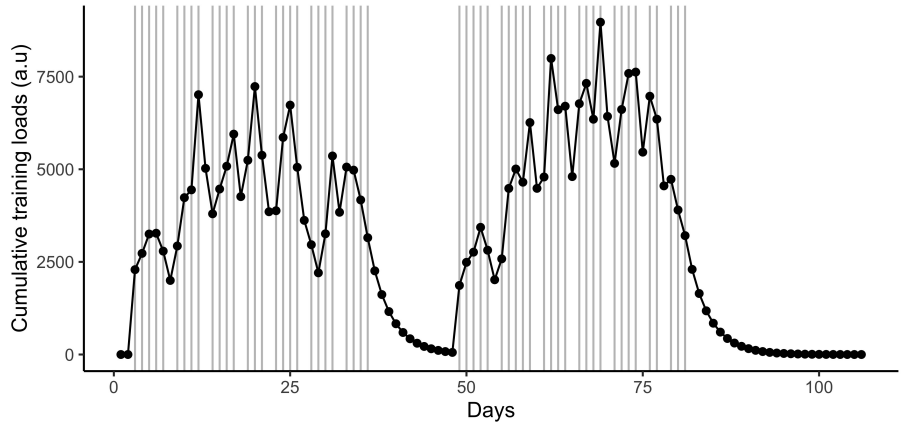
$$X_2(t) = (w * g_{\text{imp}})(t) = \sum_{l=1}^{t} w(l) \left(e^{\frac{-(t-l)}{\tau_I}} \right) \,, \quad \text{and}$$

$$X_3(t) = (w * g_{\text{ser}})(t) = \sum_{l=1}^{t} w(l) \left(\left(1 - e^{\frac{-(t-l)}{\tau_G}}\right) U \; + \; e^{\frac{-(t-TD-l)}{\tau_D}} \; | \, U - 1 \, | \right) \,,$$

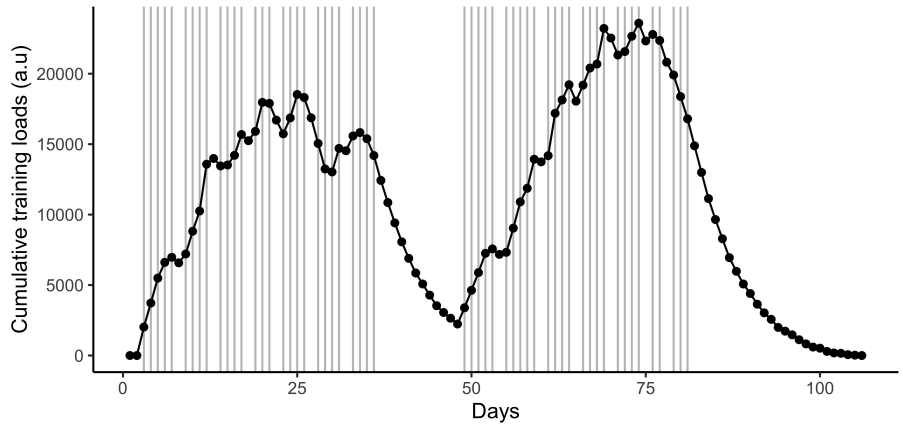
$$\text{with} \quad U = \begin{cases} 1 & \text{if} \quad t < TD \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, some characteristics components of each session were aggregated. This encompasses RPE $\{X_4, X_5\}$, averaged power $\{X_6, X_7\}$, maximal power output $\{X_8, X_9\}$, relative intensity $\{X_{10}, X_{11}\}$, session duration $\{X_{12}, X_{13}\}$ and session density $\{X_{14}, X_{15}\}$. Also, for each session ice quality $\{X_{16}\}$ and rest between two consecutive sessions $\{X_{17}\}$ were considered. Since some models may benefit from time through autocorrelated performances $y_{i,t}$, the preceding performance $y_{i,t-k}$ with $k = 1$ was included as predictor, denoted $\{X_{18}\}$. Finally, athlete $\{X_{19}\}$ was considered excepted for individually built models.

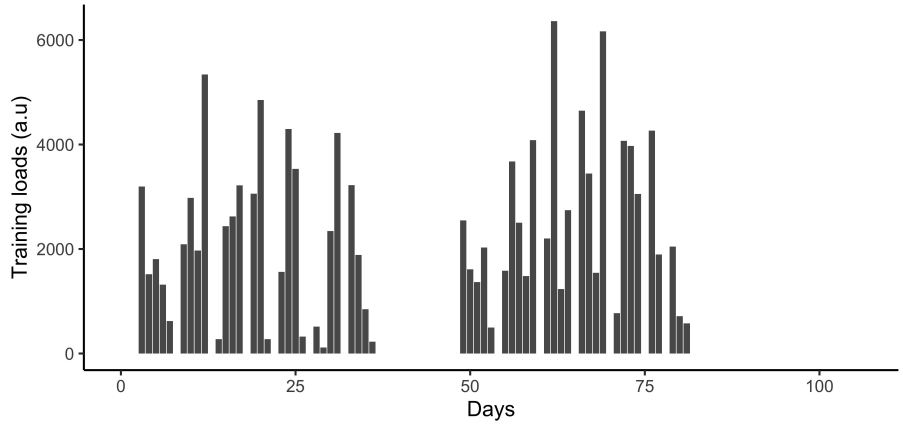
Applied to the observed data of this study a data set of $n = 248$ observations of performances associated with 19 independent variables was obtained (see Table 3.1). To formalise, allowing that $X \times Y \sim f$ with f a function of density, the built data set is a sample $S = \{(x_j, y_j)\}_{j \leq n} \sim f^n$.



(a)



(b)



(c)

Figure 3.5. – Cumulative daily training loads of a representative athlete following (a) the impulse response function (X_2, Equation 3.1) and (b) the serial bi-exponential response function (X_3, Equation 3.2). (c) illustrates the raw daily training loads X_1, expressed by w(t). In (a) and (b), dots represent daily values of the cumulative training load and vertical solid lines indicate occurrence of training sessions. Values are represented in arbitrary units (a.u).

Table 3.1. – Summary of independent variables. The two aggregation methods (impulse ans serial cumulative responses) are defined in Equations 3.1 and 3.2

Independent variables	$\mathbf{X_i}$	**Description**	**Aggregation**
Raw training load	X_1	Daily training load computed from TL_{ice}, TL_{RT}, TL_{aer}, TL_{RS}, TL_{act} (see Appendix F)	Daily recorded
Cumulative Training load	X_2, X_3	Daily computed from X_1 values	Impulse and serial cumulative responses
Rate of Perceived Exertion (RPE)	X_4, X_5	Borg category ratio (CR) 0-10 scale	Impulse and serial cumulative responses
Averaged power	X_6, X_7	On-ice sessions	Impulse and serial cumulative responses
Maximal power	X_8, X_9	On-ice sessions	Impulse and serial cumulative responses
Relative intensity	X_{10}, X_{11}	On-ice sessions (see Appendix F, Equation S1)	Impulse and serial cumulative responses
Session duration	X_{12}, X_{13}	All sessions, overall session duration	Impulse and serial cumulative responses
Session density	X_{14}, X_{15}	All sessions, effective work only	Impulse and serial cumulative responses
Ice quality	X_{16}	Subjective information quoted on a Borg 0-10 CR scale	Recorded the day of performance
Rest	X_{17}	Rest between two consecutive sessions (days)	Sum of rest days preceding the performance
Past performance	X_{18}	Significantly correlated past performance$_{t-k}$ with performance$_t$	Performance at day$_{t-k}$
Athlete	X_{19}	Athlete's id	

3.2.2. Modelling Methodology

Formally, the goal is to find a function $h : X \to Y$ which minimises the generalisation error

$$R(h) = \mathbb{P}(h(X) \neq Y) = \mathbb{E}[\mathbb{1}[h(X) \neq Y]] .$$

In practice the minimisation of R is unreachable. Instead, we get a sample set $S = (x_i, y_i)_{i \leq n} \in X \times Y$ and note the empirical error as

$$Re(h) = \frac{1}{n} \sum_{i}^{n} [\mathbb{1}[h(x_i) \neq y_i]] .$$

The objective becomes to find the best estimate $\hat{h} = \text{argmin}_{h \in \mathcal{H}} Re(h)$ with $\mathcal{H}$ the class of function that we accept to consider.

Here, four family of models are evaluated in this context. With the exception of DR, all models were individually and collectively computed (h_I and h_G, respectively).

3.2.2.1. Reference model: Variable dose-response

The time-varying linear mathematical model developed by Busso (Busso 2003) was considered as the model of reference. Formally and according to the previously introduced notation, this model is a function $h^{(busso)} : X_1 \to Y$. It describes the training effects on performance over time, $y(t)$, from the raw training loads X_1.

Training loads are convoluted to a set of transfer functions $G(t)$ and $H(t)$, relating respectively to the aptitude and to the fatigue impulse responses such as defined in Equations 1.9 and 1.10.

Combined with the basic level of performance y^* of the athlete, the modelled performance is

$$\hat{y}^{busso}(t) = \alpha_0 \, + \, k_g((w * G)(t) \, - \, ((k_h w) * H)(t) . \tag{3.3}$$

For a complete definition of the model, see Section 1.2.1.4 and Equations 3.3, 1.28, and 1.29. A non-linear minimisation of the residual sum of squares was employed in order to fit the model parameters, according to a Newton-type algorithm (Dennis Jr and Schnabel 1996).

Unlike this model of reference, the next presented models take benefit from the augmented data space $X^* = X \setminus X_1$.

3.2.2.2. Regularisation procedures

Elastic net.

In highly dimensional contexts, multivariate linear regressions may lead to unsteady models by being excessively sensitive to the expanded space of solutions.

To tackle this issue, cost functions penalising some parameters on account of correlated variables exist. On one side, Ridge penalisation reduces the space of possible functions by assigning a constraint to the parameters, thus minimising their amplitude to almost null values. On the other side, Least Absolute Shrinkage and Selection Operator (LASSO) penalisation has the capacity to fix parameters coefficient to null. The ENET regularisation combines both Ridge and LASSO penalisation (Zou and Hastie 2005). Hence, the multivariate linear model $h^{(\text{enet})} : X^* \to Y$ is

$$y_t^{(\text{enet})} = \mathbf{x_t}^t \beta + \epsilon_t \, ,$$

with $\mathbf{x} \in X^*$ the predictors, $\beta \in \mathbb{R}^d$ the parameters of the model and ϵ_t the error term. The regularisation stems from the optimisation of the objective

$$\min_{\beta \in \mathbb{R}^d} \quad \frac{1}{2} ||y_t^{(\text{enet})} - y_t||_2^2 + \lambda \left((1-\alpha)||\beta||_2^2 + \alpha ||\beta||_1 \right) ,$$

where $\alpha \in [0, 1]$ denotes the mixing parameter which defines the balance between the Ridge regularisation and the LASSO regularisation. λ denotes the impact of the penalty with $\lambda \to \infty$. For $\alpha = 0$ and $\alpha = 1$, the model will use a ridge and a lasso penalisation, respectively. Thus, for $\alpha \to 1$ and a fixed value of λ, the number of removed variables (null coefficients) increases with monotony from 0 to the LASSO most reduced model. The model was adjusted by hyper-parameters α and λ during the model selection, being part of the CV process (as described below).

Principal component regression.

Briefly presented in Section 2.2.2, PCA is of great interest in multivariate context with potential multicollinearity issues. Indeed, PCA aims to project the original data set from X^* into a new space $\tilde{X}^*$ of orthogonal dimensions (i.e PC). These dimensions are built from linear combinations of the initial variables. One may use the PC to regress the dependent variable: also known as PCR. The regularisation is performed by using as regressors only the first PC which retain the maximum of variance of the original data, by construction. In our study and according to the Kaiser's rule (Kaiser 1960), p principal components with an eigenvalue higher than 1 were retained and further used in linear regression.

Such a model, $h^{(pcr)} : \tilde{X}^* \to Y$, can be defined as a linear multivariate regression over PC as

$$y_t^{(pcr)} = \tilde{\mathbf{x}}_t^t \beta + \epsilon_t \, ,$$

with $\mathbf{x} \in \tilde{X}^* \backslash \{\tilde{X^*}_{p+1}, \ldots, \tilde{X^*}_d\}$ the predictors, $\beta \in \mathbb{R}^p$ the parameters of the model and ϵ_t the error term. In addition to being a regularisation technique by using a subset of PC only, PCR also exerts a discrete shrinkage effect on the low variance

components (the lower eigenvalue components), nullifying their contribution in the original regression model.

3.2.2.3. Random Forest

Random Forest model consists of a large number of regression trees that operate as an ensemble. RF is random in two ways, (i) each tree is based on a random subset of observations and (ii) each split within each tree is created based on a random subset of candidate variables. The overall performance of the forest is defined by the average of predictions from the individual trees (Grömping 2009). In this study, random subset of variables and number of trees were the two hyper-parameters for adjusting the model within the model selection. The model is a function $h^{(\mathrm{rf})} : X^* \to Y$.

3.2.2.4. Time series cross-validation and prediction

Since we aim at predicting daily skating performances such as non i.i.d random variables, the time dependencies have to be accounted for in the CV procedure. It ensures information from the future are not used to predict performances of the past. Hence, data were separated –respectively to the time index– into one training data set for time series CV with growing window (80 % of the total data, see Figure 3.4a) and the remaining data for an unbiased model evaluation (evaluation data set). The model selection is detailed in Algorithm 1.

Algorithm 1 Time series cross-validation

Require:
A data set of n time ordered elements, $S = \{(x_j, y_j)\}_{j \leq m}$
The number of partitions to split training data to, $K < m - s_{\min} - s_{\mathrm{val}} + 1 \in \mathbb{N}$
The minimum size of training set within a partition, $s_{\min} \geq 1$
The size of validation set within a partition, $s_{\mathrm{val}} \geq 1$
A class of functions, $\mathcal{H}$
Ensure: An optimal model h^* fitted on S
for $h^{(i)} \in \mathcal{H}$ **do**
for $k \in \{1, \ldots, K\}$ **do**
$t_{\mathrm{val}} \leftarrow s_{\min} + k$
$S_{\mathrm{train}} \leftarrow (x_t, y_t) \quad \text{with} \quad t \in [1,\ t_{\mathrm{val}} - 1]$
$S_{\mathrm{valid}} \leftarrow (x_{t'}, y_{t'}) \quad \text{with} \quad t' \in [t_{\mathrm{val}},\ t_{\mathrm{val}} + s_{\mathrm{val}} - 1]$
$h^{(i)}_{\mathrm{trained}} \leftarrow$ Train $h^{(i)}$ on S_{train}
$E[i,k] \leftarrow$Evaluate RMSE of $h^{(i)}_{\mathrm{trained}}$ on S_{valid}
end for
end for
return $h^* = h^{(i^*)}$ with $i^* = \mathrm{argmin}_i\{\frac{1}{K}\sum_{k=1}^{K} E[i,k]\}$

Algorithm 1 iteratively evaluates a class of functions $\mathcal{H}$, in which each function $h^{(i)}$ differs from its hyper-parameters values. A time ordered data set S is partitioned into training and validation subsets (S_{train} and S_{valid}, respectively). For each partition k with $k \in \{1, ..., K\}$, $h^{(i)}$ functions are fitted on the incremental S_{train} and evaluated on the fixed S_{valid} subset that occurs after the last element of S_{train}. Once $h^{(i)}$ functions are evaluated on K partitions of S, a function $h^{(i^*)}$ that provides the lowest and averaged RMSE among validation subsets defines an optimal model denoted h^*.

Model Evaluation.

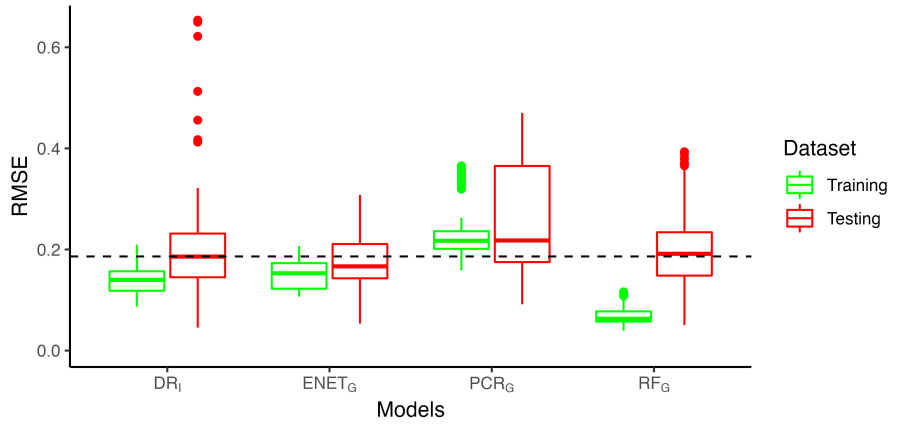
Afterwards and for each partition of S, h^* is adjusted on new time ordered training subsets S'_{train} which combines both S_{train} and S_{valid}. Then, the generalisation capability of h^* is evaluated on fixed length subsets of evaluation data S_{eval}, saved for that purpose. Note that the DR is only concerned by the model evaluation step since it has no hyper-parameters to be optimised in the model selection phase. Also, a custom-built R package has been developed for that purpose (Imbach 2020).

3.2.3. Evaluation of model performances

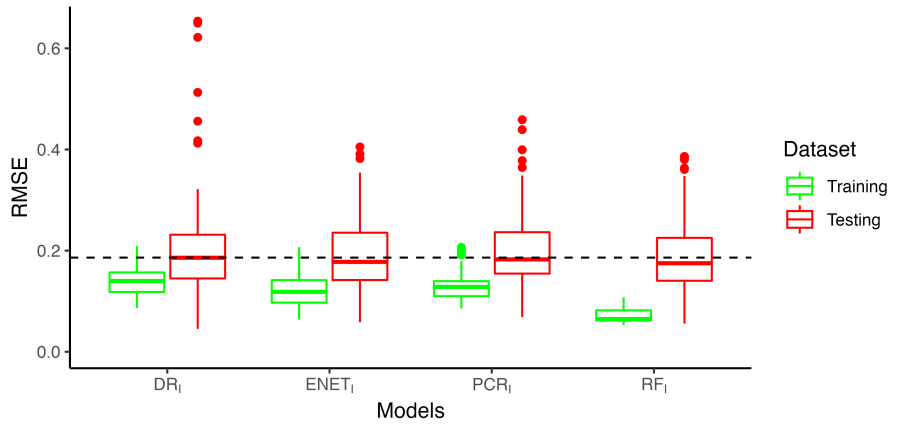
In the sequel, we sequentially evaluate the models performances in terms of generalisation capabilities and accuracy of predictions. As a reminder, each model excepted DR are built per individual and using data from the whole group.

3.2.3.1. Models generalisation

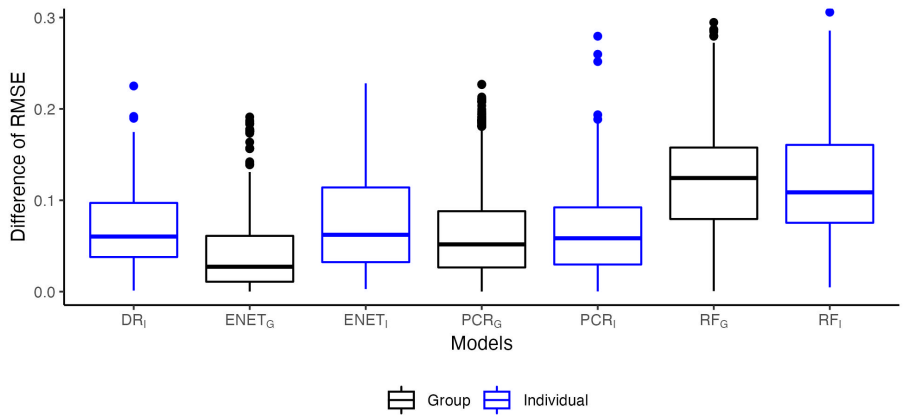
Mixed model analysis showed that both ENET and PCR models lowered the differences in terms of prediction errors between the training and evaluation data set ($\beta = -0.023 \in [-0.037, -0.007]$ 95% CI, $p = 0.003$ and $\beta = -0.057 \in [-0.065, -0.047]$ 95% CI, $p < 0.001$ for individually-computed elastic net regularisation ($ENET_I$) and group-computed elastic net regularisation ($ENET_G$); $\beta = -0.026 \in [-0.040, -0.011]$ 95% CI, $p < 0.001$ and $\beta = -0.032 \in [-0.041, -0.022]$ 95% CI, $p < 0.001$ for indivual-computed principal component regression (PCR_I) and group-computed principal component regression (PCR_G), respectively). A significant effect of the model class on the generalisation risk was also reported ($p < 0.001$, $\eta^2 = 0.23 \in [0.20, 0.26]$ 95% CI). The most generalisable models were ENET and PCR models computed on overall data, followed by individual based models. Generally, group-built models likely provided a greater generalisation capability than individual based models ($\beta_{diff} = -0.0144$, $p < 0.001$, $\eta^2 = 0.01 \in [0.00, 0.01]$ 95% CI). A summary of model pairwise comparisons is provided in Table 3.2.



(a)



(b)



(c)

Figure 3.6. – Distributions of models' performance. (a) shows RMSE distributions of each individual models and (b) the models computed on the whole group. Within boxplot midline represents the median of the distribution. All of them are compared to the dose-response (DR) model. (c) represents the differences of RMSE between training and testing data for each model.

Table 3.2. – Summary of models pairwise comparisons for generalisation and prediction abilities. β_{diff} represents the marginal mean difference of the RMSE distribution between the DR model and its comparison.

Comparison	β_{diff}	t ratio	p	Criterion
$DR_I - ENET_G$	0.057	-11.841	< 0.001	Generalisation
$DR_I - PCR_G$	0.032	6.644	< 0.001	Generalisation
$DR_I - PCR_I$	0.026	3.365	0.004	Generalisation
$DR_I - ENET_I$	0.023	2.933	0.018	Generalisation
$DR_I - RF_G$	-0.027	-5.649	< 0.001	Generalisation
$DR_I - RF_I$	-0.028	-3.831	< 0.001	Generalisation
$DR_I - ENET_G$	0.041	5.607	< 0.001	Prediction
$DR_I - RF_G$	0.022	3.067	0.012	Prediction
$DR_I - ENET_I$	0.021	2.112	0.156	Prediction
$DR_I - RF_I$	0.018	1.789	0.294	Prediction
$DR_I - PCR_I$	0.016	1.537	0.438	Prediction
$DR_I - PCR_G$	-0.042	-5.779	< 0.001	Prediction

Cross-validation outcomes revealed significant heterogeneity in performances of models, even though the differences remain small regarding the total time of skating trials (see Table 3.3). The main criterion of interest, generalisation, was significantly greater for both ENET and PCR models than the DR model. One can explain this result by the capabilities of the statistical models to better catch the underlying skating performance process using up to 19 independent variables when associated with regularisation procedures. Conversely, the DR model relies on two antagonistic components strictly based on the training load dynamics. It does not deal with any other factors that may greatly impact the performance (e.g. psychological, nutritional, environmental, training-specific factors) (Avalos, Hellard, and Chatard 2003; Mujika, Busso, Lacoste, et al. 1996; Stone, Stone, and Sands 2007). Thus, such a conceptual limit can be overtaken by employing multivariate modelling that may result in greater comprehension of the training load - performance relationship for the purpose of future predictions (Avalos, Hellard, and Chatard 2003; Hellard, Avalos, Lacoste, et al. 2006). As presented in Section 1.2.1.1, Piatrikova, Willsmer, Altini, et al. 2021 extended the former FFM framework (Banister, Calvert, Savage, et al. 1975) to account for some psychometric variables as model inputs. Despite the fact that the authors reported improved goodness of fit for this multivariate alternative, attributing impulse responses to these variables might question the conceptual framework behind the model.

Distributions of RMSE from training and testing data sets allow us to establish a generalisation model ranking (see Table 3.2). Regularised linear models computed on overall data offered a better generalisation. This finding is essential because by handling the bias-variance trade-off, models are more suited for capturing a proper underlying function that maps inputs to the target even on unknown data.

Hence, it allows further physiological and practical interpretations from the models such as the remodelling process of skeletal muscle involved by exercise, dynamically represented by exponential growth and decay functions (Philippe, Borrani, Sanchez, et al. 2019). The sample size might partly explain this result. It is clear that statistical inference on small samples leads to bad estimates and consequently to bad performances in prediction (Cui and Gong 2018; Kelley and Maxwell 2003). A greater sample size obtained by combining individual data led to more accurate parameter estimates, being more suitable for sports performance modelling (Avalos, Hellard, and Chatard 2003). That is particularly important to consider when we aim to predict a very few discipline-specific performances throughout a season. However, predicting non-invasive physical quality assessments which can be daily performed (e.g. squat jumps and their variations for an indirect assessment of neuromuscular readiness (Watkins, Barillas, Wong, et al. 2017), short sprints) may be an alternative for small sample size issues. In our case, standing start time trials over 1.5 laps allowed for the coach to evaluate underlying physical abilities of the skating performance several times a week. Also, regularisation tends to stabilise parameters estimators and favour the generalisation of the models. For instance, multicollinearity may occur in high-dimensional problems. Stochastic models generally suffer from such conditioning. One would note that the ENET and PCR models attempt to overcome these issues in their own way by (i) penalising or removing features – or both – that are mostly linearly correlated and (ii) by projecting the initial data space onto a reduced space, which is optimised to keep the maximum of variance of the data from linear combinations of the initial features. Both approaches limit the number of unnecessary – or noisy – dimensions. In contrast, in this study, non-linear machine learning models (group-computed random forest model (RF_G) and indivual-computed random forest model (RF_I)) expressed a lower generalisation capability than linear models even when models combine data from several athletes. Such models may be powerful in multidimensional modelling but require an adequate data set, particularly ones with sufficient sample size. Otherwise, model overfitting may occur at the expense of inaccurate predictions on unknown data.

3.2.3.2. Accuracy of predictions

Root mean square errors reported on evaluation data using mixed model analysis indicated that $ENET_G$ was the most contributing model in lowering the prediction errors ($\beta = -0.041 \in [-0.055, -0.027]\ 95\%\ CI$, $p < 0.001$), followed by RF_G as shown in Table 3.2. Accordingly, a significant model class effect on prediction errors was reported ($p < 0.001$, $\eta^2 = 0.18 \in [0.15, 0.21]\ 95\%\ CI$). Computing models over a larger population (i.e. group-based models) showed only a trend in favour of group-based models over the errors response rate ($p = 0.146$).

Distributions of RMSE on data used for model evaluation have shown heterogeneous variance between models. The greatest standard deviations were found

for individually-computed variable dose-response model (DR_I) and PCR_G with $\sigma = 0.053$ and $\sigma = 0.062$ respectively. The ENET, PCR_I and RF models provided more consistent performances with lower standard deviations comprised within $[0.023; 0.027]$ and $[0.012; 0.017]$ intervals for individual and group computed models, respectively. Note that the greatest errors on evaluation data were systematically attributed to one particular athlete. In average, predictions made from this athlete led to greater RMSE than ones made from other athletes ($p < 0.001$, $\beta_{diff} = 0.22[0.163, 0.286]95\%CI$). Mean values of coefficient of determination (R^2) indicated that weak linear relationships between performance and predictors were identified by models ($R^2 \in [0.150; 0.206]$). The highest averaged R^2 value but also the greatest standard deviations were reported for DR_I models ($R^2 = 0.206 \pm 0.093$). However, significant differences of averaged R^2 were only found for $ENET_I$, RF_G and PCR_G ($\beta = -0.056$ $[-0.10; -0.01]$ 95% CI, $p = 0.02$; $\beta = -0.041$ $[-0.08; -0.01]$ 95% CI, $p = 0.02$ and $\beta = -0.036$ $[-0.07; -0.01]$ 95% CI, $p = 0.04$ respectively). A summary of model performances is provided in Table 3.3.

Predictions made from the two most generalisable models – $ENET_G$ and PCR_G – and the reference DR_I illustrate the sensitivity of models for a representative athlete (Figure 3.7). Performances modelled from DR_I model were relatively steady and less sensitive to real performance variations. Standard deviation calculated on data used for model evaluation supported such a smooth prediction with $\sigma = 0.015$, $\sigma = 0.071$ and $\sigma = 0.062$ for DR_I, PCR_G and $ENET_G$, respectively. Regarding $ENET_G$, the greatest standardised coefficients were attributed to the auto-regressive component (i.e. the past performance) such as $\beta = 0.469$, followed by the athlete factor and then impulse and serial bi-exponential aggregations. For regression, PCR_G used the three first principal components explaining 52.3%, 16.5% and 7.6% of the total variance, respectively. Details about models' parameters as well as principal component compositions are available in Appendix G.

Table 3.3. – Summary of the predictive models. According to model families, criteria were averaged among folders and displayed with their standard deviation. For individual models, averaged values of hyper parameters are displayed along with lower and upper recorded values. The greatest performance among criteria is listed in bold type.* indicates the DR_I as the reference model and specification of its averaged parameters.

Model	R^2	MAE	RMSE	Hyper parameters*
DR_I^*	**0.206 ± 0.093**	0.189 ± 0.055	0.225 ± 0.053	$\overline{k_1}$ = -3.95e-05, $k_1 \in$ [-4.85e-05; -3.19e-05]
				$\overline{k_3}$ = -7.75e-09, $k_3 \in$ [-4.01e-09; -1.71e-08]
				$\overline{\tau_1} = 36.02, \tau_1 \in [25.82; 42.28]$, $\overline{\tau_2} = 22.57, \tau_2 \in [14.58; 26]$, $\overline{\tau_3} = 5.23, \tau_3 \in [4.33; 6.67]$
$ENET_I$	0.150 ± 0.010	0.169 ± 0.020	0.197 ± 0.023	$\overline{\alpha} = 0.176, \alpha \in [0; 0.6]$, $\overline{\lambda} = 0.273, \lambda \in [0; 1]$
PCR_I	0.164 ± 0.068	0.173 ± 0.025	0.201 ± 0.027	$\overline{n\,comp} = 1.918, n\,comp \in [1; 3]$
RF_I	0.193 ± 0.074	0.170 ± 0.023	0.199 ± 0.024	$\overline{mtry} = 8.90, mtry \in [1; 17]$
$ENET_G$	0.179 ± 0.063	**0.150 ± 0.010**	**0.176 ± 0.012**	$\alpha = 0.28, \lambda = 0.02$
PCR_G	0.17 ± 0.053	0.22 ± 0.044	0.259 ± 0.062	$ncomp = 3$
RF_G	0.164 ± 0.069	0.163 ± 0.017	0.195 ± 0.017	$mtry = 16$

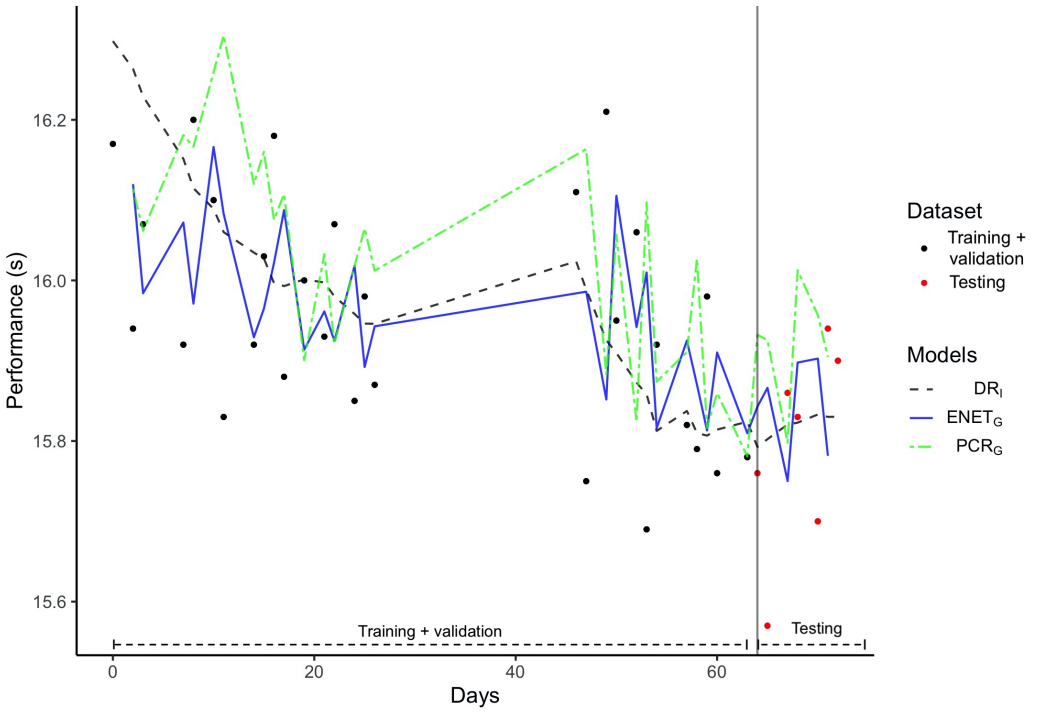


Figure 3.7. – Modelled performance of a representative subject. Solid and dashed lines represent the DR model and the two models offering the best generalisation. On this example, the training data set (80% of the data that combines training and validation subsets) and testing data set (20% of the data, the testing subset) areas are separated by the vertical solid line. Fitted parameters of the DR model were k_1 = -2.45e-05, $k3$ = -2.58e-09, $\tau_1 = 39$, $\tau_2 = 26$, $\tau_3 = 5$. Hyper-parameters of the PCR and ENET models were $n\ comp = 3$ and $\alpha = 0.28$, $\lambda = 0.02$. Note that in accordance with ENET and PCR models, only day of performances are displayed through DR.

As reported previously and with the exception of PCR_G, models were more accurate in prediction than DR (Table 3.3). The large averaged RMSE, as well as large standard deviations provided by the DR among performance criteria, tend to agree with the literature since the model is prone to suffer from weak stability and ill-conditioning raised by noisy data that impact its predictive accuracy (Hellard, Avalos, Lacoste, et al. 2006; Ludwig, Asteroth, Rasche, et al. 2019). It evokes that linear relationships between the two components "Aptitude" – "Fatigue" and the performance are unclear. However, because of a lack of cross-validation procedures on impulse response models and particularly the DR model employed in our study, our results cannot be validly compared with the literature. Despite lower standard deviations of R^2 reported by ENET and PCR models, the weak averaged R^2 values suggest that linear models can only explain a few part of the total variance. Note

that all linear models are concerned (including theDR), since the differences in averaged R^2 between models are relatively small and only significant for $ENET_I$, RF_G and PCR_G models. Therefore and if the data allow it (i.e. a sufficient sample size and robustly collected data), non-linear models may still be effective and should be considered during the modelling process.

The sensitivity of models according to gains and losses of performances differed between the two most generalisable models – $ENET_G$ and PCR_G – and the reference DR (see Figure 3.7). Such differences can be explained by the influence of variables that may affect performance, other than training loads dynamic (e.g. ice quality the day of performance, cumulative training loads following a serial and bi-exponential function, the last known performance) or a DR model failure in parameter estimates regarding to the variability of the data. Indeed, parameters estimates of $ENET_G$ supported that since changes in skating performance were mostly explained through the past performance, weighted by individual properties and to a lesser degree by training-related parameters. The PCR_G used a different approach for the same purpose and greatly relied on training related aggregations as well as environmental and training programming variables (see Appendix G). However, this applied example does not inform us about neither the generalisation ability of models nor accuracy of predictions because it concerns only a particular set of data, where the selected models (i.e. with optimal hyper-parameters) are trained on the first 80 % of data and evaluated on the 20 % remaining data. In addition, since model estimates greatly depend on the sample size, we might expect significantly different estimates with more data.

3.2.3.3. Limits and conclusion

This study presents some limits but reveals some major methodological considerations. The first limit concerns the data we used and particularly the criterion of performance: standing start time trials few times a week during an approximately 3-months period. Even though being a very discipline-specific test in which athletes are familiar and being conducted in standardised conditions, each test requires high levels of arousal, buy-in, motivation and technique. Therefore, psychological states and cognitive functions monitoring such as motivation and attentional focus (Gillet, Berjot, Vallerand, et al. 2012; Ille, Selin, Do, et al. 2013) should have been done prior to performing each trial. A concrete example is provided through the Figure 3.7, where $ENET_G$ greatly penalised the training correlated features and kept the influence of the auto-regressive component predominant over other features. It might either be the consequence of an inference issue due to the relatively small sample size or a lack of informative value of training related features that do not allow to explain changes in skating performance. Also, both reasons support models failure in predicting skating performances of one particular athlete, who showed significant greater errors of prediction. It emphasises the importance of measuring the "right" variables for performance modelling purposes, particularly if

the sport-specific performance involves various determining factors.

Secondly, the time series CV presented here has a particular cost, most notably when only a few data are available (e.g. when models are individually computed). The rolling origin re-calibration evaluation performed as described by Bergmeir and Benı+tez 2012 implies a model training only on an incremental sub-sequence of training data. Hence, the downsized sample size of the first training sub-sequences might cause model failure in parameter estimates and, consequently, an increase of prediction errors.

Then, training and testing data sets present some dependencies. In order to evaluate models on fully independent data, some modifications of the current CV framework exist at the expense of withdrawing even more data in the training procedure. As presented in Figure 3.4c and according to Racine 2000, *hv - block cross-validation* would be one of the least costly alternative to the CV used in our study, requiring a certain gap between each training and validation subsets. However, due to limited sample size, we voluntary chose not to adapt the original CV framework described in Algorithm 1. Nonetheless, such an alternative would be recommended for researchers and practitioners to predict athletic performances with significant dependencies and sufficient sample size.

Finally, *backtesting* was performed in order to evaluate model performances on historical data. From a practical point of view, models can predict the future performance following a given feature of data known until day t. However, the contribution of training load responses modelling also concerns training after-effects simulations over a longer time frame. Having identified a suitable model, practitioners may pinpoint key performance indicators –specific to the discipline of interest– and confront model estimates to field observations. Then, simulations of these independent variables within their distributions would allow practitioners and coaches to simulate changes in performance following objective and subjective measures of training loads, and any performance factors that are monitored. Conditional simulations that consider known relationships between independent variables (e.g. relationships between training load parameters) (Casamichana, Castellano, Calleja-Gonzalez, et al. 2013; Noble, Borg, Jacobs, et al. 1983) may improve the credibility of simulations.

In summary, a model selection according to a time-series CV yielded to slight but significant greater generalisation capability of ENET and PCR models than DR. Indeed, ENET showed the greatest performances both in terms of generalisation and accuracy in prediction. In addition, increasing the sample size by computing models on the whole group of athletes led to more performing models than the individually computed ones. This is in accordance with the singularity of training responses exposed so far since athlete identification is considered as a predictor in the models. The methodology provided in this section can be reemployed in other contexts and sports to optimise elite sport performance through training protocols simulations. Finally, model evaluation according to a robust method would be a requisite for any physiological and practical interpretation to make future athletic

performance predictions.

The results presented so far showed how much performances can vary among a few classes of models, according to the generalisation capabilities and model accuracy. However, several models could be employed for that purpose, still under multicollinearity and high dimensionality problems. In addition, DR seemed to greatly suffer from its univariate configuration despite being built on a sound basis. It leaves a place for extending its existing framework either by increasing its complexity to a higher dimensionality or by integrating it into elaborated ML frameworks.

Take home message.

— Multivariate linear approaches within regularisation procedures provided the greatest performances in terms of generalisation and accuracy. Also, multivariate models took over univariate models, in particular when the performance relies on several factors rather than solely the TL dynamics.
— According to our data set, models computed using the whole available data also provided better generalisation performances.
— Since time-dependencies usually exists in athletic performance modelling problems, it has to be accounted for in any model selection procedure.

3.3. Fitness-fatigue models: advantages, conceptual issues and contribution from machine-learning

As presented in Section 1.2.1, FFMs related to a collection of impulse response models that were first developed for athletic performance description purposes and then for optimising training protocols. The latter objective generally came with simulations of daily TL for finding an optimal training sequence (*i.e.* defined by the amount of TL and training occurrence within a time window) that would theoretically maximise an athletic performance (Clarke and Skiba 2013; Méline, Mathieu, Borrani, et al. 2019; Sanchez, Galbès, Fabre-Guery, et al. 2013; Thomas and Busso 2005; Thomas, Mujika, and Busso 2008, 2009). Hence, physiological and practical interpretation would arise for the benefit of technical and medical teams around athletes. Although the framework behind FFMs seems promising, it might be improved for better predictive performances purposes.

3.3.1. Making the most of control theory

Using the formalism of transfer functions as a model of relationships between training doses and fitness and fatigue states provide several advantages. First,

while Banister, Calvert, Savage, et al. 1975 consider the human performance, such as the result of the difference between two simple first-order transfer functions, the model could be extended to more complex transfer functions. It would allow modelling much more sophisticated dynamic relationships between exercise and state variables, according to the complexity and interactions between physiological processes involved in human (Lambert, Gibson, and Noakes 2005). The only drawback to this extension would be the loss of direct physiological interpretation, but the model identification phase would not be further complicated.

Secondly, transfer functions are the basic tools of control theory (Sontag 2013). This branch of mathematical optimisation deals with finding a control for a dynamic system over a period of time, such that an objective function is optimised. In classical linear quadratic (LQ) optimal control problems, the resulting control law (i.e. the training doses here) can be analytically provided from the algebraic structure of the dynamic system and the expected output over a finite temporal horizon. Hence, the optimal control law is a time-varying linear function of the state variables. Control theory framework thus allows for identifying the optimal training doses in a program, in order to reach a given performance setpoint while minimising some energy criterion that can be seen as a weighted function over state variables (e.g. fitness and fatigue) and inputs (e.g. daily training doses) in LQ problems. It is a matter of compromise between energetic cost and response time of the closed-loop system, and depends on the decision-maker policy (e.g. the coach's strategy). Although the first FFMs appeared more than forty years ago, their use to design the optimal training programming is essentially envisaged through simulations. In contrast, their main advantage relies on their algebraic structure for control purposes.

Finally, the unexploited algebraic structure of FFMs would also provide state observers. By definition, a state observer is a system that provides an estimate of the internal state of a given real system from measurements of the input and output of the real system (Luenberger 1971). In our field of application, it could be used to precisely estimate the state variables of athletes that are considered in the model (i.e. fitness and fatigue for FFMs) or to adjust the model through performances' observations.

3.3.2. Current Fitness-fatigue framework and related issues

The input: quantification of training.

The first step of any training effect modelling using FFMs requires quantifying the training itself. The models being mostly applied to individual and particularly endurance sports, a few methods were used to quantify the training dose as presented in Section 1.1. Hence, the model input defined by a discrete function $\omega(i)$ can take various expressions, be the TL being objectively or subjectively quantified.

On this basis, the training sessions are the only cause of adaptations. That means training responses are independent of any other external factors to training,

yet known to impact athletic performance but not being accounted for in the model (e.g. environmental factors, nutritional and psychological status). Hence, two identical training sessions that occur at different training stages would induce similar adaptations and responses. Besides, different training sessions (e.g. a LI and prolonged exercise, and a HI and short exercise) may result in similar TL estimates (and so fitness and fatigue states) since training related parameters are resumed into a single variable (Busso and Thomas 2006), despite specific responses and adaptations to exercise exist (Fry and Kraemer 1997; Helgerud, Høydal, Wang, et al. 2007). Finally, athletes usually practice endurance and resistance training, and other disciplines to enhance performance. Since FFMs are sensitive to the nature of the model input (Vermeire, Van de Casteele, Gosseries, et al. 2021), a consistent training quantification method that is not biased by the type of training is required across training sessions.

Taking this stand, a univariate configuration of FFMs reduces the space of dimensions around adaptations to training into one single dimension, solely characterised by the training quantification. This is at the expense of all relevant information that can be caught and may explain a part of athletic performance, even if the training quantification is objectively well estimated. It also questions training quantification based on arbitrary methods, which might bring "noise" in the modelling in case of a rough appreciation of the exercise demand by the coach.

The function: a physiological approximation.

Attempting to model athletic performance upon a mathematical representation of physiological principles is commendable. However, it implies being confident in the model itself, leaving no room for vague theoretical approximations. Among the aforementioned variants of the original FFM, improving model complexity (e.g. by adding components in the model) do not guarantee the best model performances (see Section 1.2.1) (Busso, Carasso, and Lacour 1991), even though such models are supposed to better represent the physiological responses. Therefore, the pertinence of adding components to the most basic structure (i.e. only based on the fitness component) and, more generally, the theoretical hypothesis behind FFMs might be questioned.

The output: the performance.

Finally, FFMs attempted to model either an athletic performance during a competitive season, a physical ability that relates to athletic performance (e.g. mean power or velocity sustained on shorter distances than ones performed during competitions) (Busso 2003; Millet, Candau, Barbier, et al. 2002) or a physiological state (Chalencon, Pichot, Roche, et al. 2015; Williams, West, Howells, et al. 2018). In general, choosing the appropriate model output has a strong implication in the modelling process. Modelling changes in physical ability instead of a full discipline-specific performance may allow for repeating less invasive and better-controlled testing all along a training process. However, whatever form the output takes (i.e.

an athletic performance or a physiological state), its multifaceted nature involves factors that are not considered in the univariate model. Therefore, the training history merely characterised by training loads may only explain a part of output changes, resulting in a lack of model performances.

To summarise, FFMs ability to predict changes in athletic performance is greatly impacted by univariate modelling issues and a simplification of human physiological adaptations to exercise and training. Moreover, considering only the training loads responsible for athletic performance changes implies neglecting all related confounding variables that influence both explanatory and target variables, causing spurious associations between input and output of the model.

3.3.3. A machine-learning perspective of the problem

As mentioned in Section 1.2.3, statistical and ML models with a different approach to the problem. Attempting to predict target variables from sets of co-variables, they foster multivariate modelling that comprises TL variables and all measured variables that may explain changes in athletic performance. In addition, they allow for modelling non-linear relationships between training-related parameters and the target, making them greatly attractive in particular when compared to FFMs in sports (Mitchell, Rattray, Fowlie, et al. 2020). However, comparing performances of ML models to FFMs, this is not surprising because the latter represents a restrictive class of models based on strong assumptions (e.g. impulse nature of the response to exercise, the athletic performance resulting of the difference between two fitness and fatigue states), which are essentially linear. Such a comparison is also largely biased by the higher degree of freedom of ML models. Therefore, we believe that ML models should not be considered as an alternative to FFMs, but a way to improve and broaden FFMs applications instead.

Expert knowledge and strong physiological assumptions that led to the mathematical framework of FFMs represent valuable information that could be used inside ML models. In addition, raw data may also be considered to keep the maximum of information and thus advance the athletic performance modelling through an inclusive perspective. Nevertheless, no studies combining FFMs and ML models have been carried out.

Despite several extensions of the former FFM (Banister, Calvert, Savage, et al. 1975) have been developed for predictive applications, there is no consensus about the optimal mathematical structure to be retained. Each of the FFMs variants have their advantages and drawbacks, but they remain close in terms of predictive performances while being heterogeneous in terms of complexity. In addition, predictions made from these models suffer from high bias and low variance, particularly when the target is greatly sensitive to other variables than training load dynamics. Initiated by Wolpert 1992, a stacking algorithm from ensemble learning methods could be used to increase FFMs predictive performances. Stacking let us consider a set of FFMs as base-models for predicting athletic performances through

fitness and fatigue states, along with proven ML models that include any other variables of interest, of any kind. Predictions of all base-models (FFMs and ML models) are aggregated through a "meta-regressor" such as a regularised linear regression. The overall process is presented in Figure 3.8.

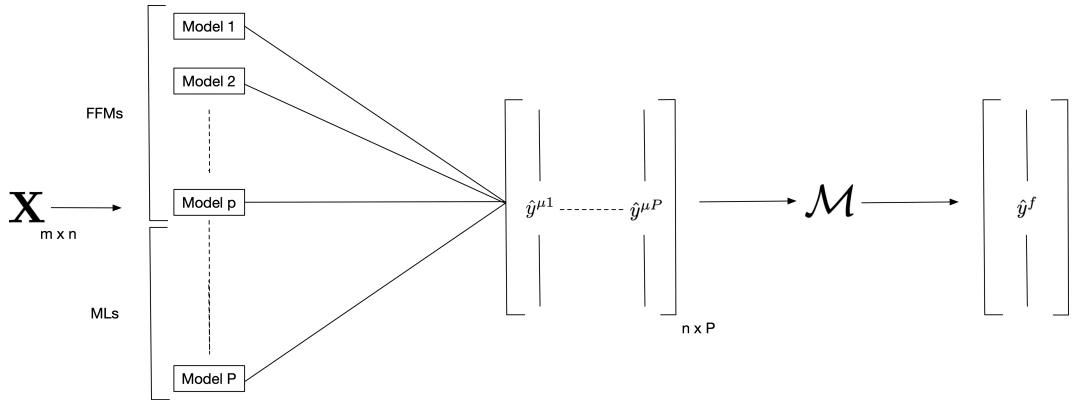


Figure 3.8. – Stacking ensemble learning using several fitness-fatigue and ML models.

Briefly, let $\mathbf{X}_{\text{m x n}}$ being a first level training data set of m features and n observations. Base-models are composed of various FFMs and ML models which are concurrently trained (within validation procedures such as CV, see Section 3.1.2). Predictions made from the models constitute an out-of-sample prediction matrix (i.e. a second level training data set) of dimension n x P base-models. Finally, a combiner –or a meta-model– denoted $\mathcal{M}$ is trained on these data to predict the final outcome $\hat{y}^f$.

To summarise, the meta-model could thus be used to find the best combination of FFMs and ML models for better prediction purposes Witten and Frank 2002. In addition, opting for inherently interpretable "white box" models (i.e. models which provide understandable mappings between inputs and outputs through closed formulas or graphs, such as linear regressions or decision trees) as meta-models Loyola-Gonzalez 2019; Rudin 2019 could improve experts understanding of the FFMs shortcomings and how ML models can compensate them. In addition, stacked ensembles do not require a larger sample size than if the models were used separately.

3.3.4. Take home message

— Fitness-Fatigue models rely on expert knowledge and could be extended to more complex functions, including other athletic performance factors for prediction purposes.
— Through ensemble learning methods such as stacking, machine-learning approaches are not alternatives to Fitness-Fatigue models but rather a way of

improving their predictive capability while preserving expert information in the modelling.

3.4. Using global positioning systems for modelling athletic performances

In Section 3.2 we modelled an athletic performance in regards to models generalisation and using data from elite short-track speed skaters. From observable parameters such as the time and distance covered that are measured using basic systems, we derived some mechanical parameters to form training-related parameters and as well as a TL index. However, the technological improvement usually ascribed to team sports allows for measuring the least effort through several ways, with high precision, and in ecological conditions (see Section 1.1.1.3). Therefore, practitioners have to deal with a large amount of data that may stem from various sources, in which the explanatory capacities remain to be investigated.

Global positioning systems are gold standard systems of measurements in sports field. Widely used for monitoring purposes (Akenhead and Nassis 2016; Bourdon, Cardinale, Murray, et al. 2017; Buchheit, Al Haddad, Simpson, et al. 2014; Cardinale and Varley 2017; Coppalle, Rave, Ben Abderrahman, et al. 2019; Jennings, Cormack, Coutts, et al. 2010; Kupperman and Hertel 2020; Malone, Lovell, Varley, et al. 2017; Ravé, Granacher, Boullosa, et al. 2020; Ryan, Kempton, and Coutts 2020; Theodoropoulos, Bettle, and Kosy 2020), they allow quantifying the exercise demand from both GPS and IMU signals (Gómez-Carmona, Bastida-Castillo, Ibáñez, et al. 2020) with an increasing accuracy (see Section 1.1.1.3 for systems validity). In addition, the growing interest in injury prevention makes researchers focus also on injury prediction problems from GPS-derived insights (Claudino, Oliveira Capanema, Souza, et al. 2019; Malone, Lovell, Varley, et al. 2017; Maupin, Schram, Canetti, et al. 2020; Rossi, Pappalardo, Cintia, et al. 2018; Vallance, Sutton-Charani, Imoussaten, et al. 2020).

The aggregated features returned to the customer are defined by the manufacturer, usually returning a set of fifteen essential variables which are derived from the GPS and accelerometer measurements. The common pooling methods used for extracting features from GPS raw data usually rely on means and sums of the measurements or their derivatives (*e.g.* a total distance run at a specific speed interval or an averaged pace within defined intervals, be they empirically fixed or individually set). The simplest features being relevant in practical terms, raw GPS data might provide some other features of interest, such as ones supported by physiological meaning (Osgnach, Poser, Bernardini, et al. 2010; Osgnach and Prampero 2018; Prampero and Osgnach 2018). Otherwise, one should consider any other features but it requires to have access to the raw data.

Using GPS data for predicting athletic performances remains challenging. First, defining an athletic performance such as a *physical* performance in which interactions with opponent and environment are lowered. It usually necessitates programming specific testing sessions in controlled conditions through the season. Even though being necessary for any performance monitoring, multiplying testing sessions involve

time and investigation costs, injury exposure, psychological state disruptions, and adjustments of training plans. However, Morin, Le Mat, Osgnach, et al. 2021 recently proposed a promising method for assessing a player performance while practising football without performing any specific tests. In brief, the method consists of modelling individual acceleration - velocity profiles from GPS raw data. Such profiles have practical meanings, notably for monitoring changes in athletic properties (by analogy to F-V profiles). However, it remains to be further validated for athletic performance and injury predictions purposes.

From these bases and according to the literature, three research issues emerge:

1. The value of the common aggregated features that are delivered by GPS sensors manufacturers regarding predictive applications.
2. The predictability of such acceleration – velocity profile using only GPS related features.
3. The use of raw GPS data for extracting new valuable features for the same purposes.

In the sequel, we investigate these issues using data from an elite football team.

3.4.1. Methodological approach

3.4.1.1. Population studied

Data from the *FC Lucerne* football club were collected over a 1.5 season period (2019–2021). The team evolves in the Superleague division, the highest division in Swiss professional football. A total of 196 training sessions and 74 games were stored in a cloud-hosted multi-model database (ArangoDB, CA, USA). For each session, raw GPS data (Fieldwiz V1, CH) and summarised features (see Table 3.4 for details) of each player were stored in the database as .json files. A total of 42 players were initially recorded, including regular professional players and young hopes.

3.4.1.2. Data set definition

Independent variables.

Independent variables are summarised in Table 3.4. Let $\mathcal{X} \subset \mathbb{R}^d$ with $d \in \mathbb{N}$ be the domain of definition of the random variable $X = \{x_1, \ldots, x_d\}$. The variable X is thus defined as a vector of d dimensions, composed of aggregations of the summarised features given by the GPS software. Aggregated features can take two forms:

1. The average of summarised features $\bar{X} = \{\bar{x_1}, \ldots, \bar{x_d}\}$ (see Equation 3.4).
2. An exponential weight according to a *softmax* function (see Equations 3.5 and 3.6).

$$\bar{X}_i = \frac{1}{N} \sum_{j=1}^{N} X_{i,j} \quad , 1 \leq i \leq d \tag{3.4}$$

and

$$X_d = \sum_j w_j X_{d,j} \tag{3.5}$$

with

$$w_j = \sigma(\mathbf{t})_j^{\beta} = \frac{e^{-\beta t_j}}{\sum_{k=1}^{K} e^{-\beta t_j}} \quad \forall j \in \{1, \ldots, K\} \,. \tag{3.6}$$

In Equation 3.5, X_d denotes an aggregated feature weighted by a scaling factor w_j. w_j is determined by a *softmax* function $\sigma(\mathbf{t})_j^{\beta}$ in which $\mathbf{t}$ is a time vector describing the distance of events to the game of interest and β denotes a scale parameter that sets the sensibility of the exponential decay weighting function.

For both aggregation methods, we arbitrarily set a window L of size $L = 5$. It refers to the summarised predictor sets given by the GPS software that are pooled according to the last L sessions (either practice or game) preceding the game of interest. Since the frequencies of sessions are heterogeneous, the number of days preceding the game to be predicted may differ over the weeks.

Table 3.4. – Summary of independent variables

Independent variables	X_i	Description
TD	X_1	Total distance covered by a player (km)
D_{0-5}	X_2	Distance covered at a running speed below 5 $km.h^{-1}$
D_{5-10}	X_3	Distance covered at a running speed between 5 $km.h^{-1}$ and 10 $km.h^{-1}$
D_{10-15}	X_4	Distance covered at a running speed between 10 $km.h^{-1}$ and 15 $km.h^{-1}$
D_{15-21}	X_5	Distance covered at a running speed between 15 $km.h^{-1}$ and 21 $km.h^{-1}$
D_{21-24}	X_6	Distance covered at a running speed between 21 $km.h^{-1}$ and 24 $km.h^{-1}$
D_{24-30}	X_7	Distance covered at a running speed between 24 $km.h^{-1}$ and 30 $km.h^{-1}$
D_{30}	X_8	Distance covered at a running speed over 30 $km.h^{-1}$
HID_15	X_9	Distance covered at high intensity (over 30 $km.h^{-1}$)
HID_21	X_{10}	Distance covered at high intensity (over 21 $km.h^{-1}$)
Acc_3	X_{11}	Number of accelerations performed over 3 $m.s^{-2}$
$Acc_{3.5}$	X_{12}	Number of accelerations performed over 3.5 $m.s^{-2}$
Acc_4	X_{13}	Number of accelerations performed over 4 $m.s^{-2}$
Dec_3	X_{14}	Number of decelerations performed over 3 $m.s^{-2}$
$Dec_{3.5}$	X_{15}	Number of decelerations performed over 3.5 $m.s^{-2}$
Dec_4	X_{16}	Number of decelerations performed over 4 $m.s^{-2}$
$Sprint_{24}$	X_{17}	Number of running sprints over 24 $km.h^{-1}$
$Speed_{avg}$	X_{18}	Averaged running speed of the session ($km.h^{-1}$)
$Speed_{max}$	X_{19}	Maximal running speed reached during the session ($km.h^{-1}$)
Acc_{max}	X_{20}	Maximal accelerations performed during the session ($m.s^{-2}$)
Player	X_{22}	Player's id

Dependent variables.

In order to investigate the effect of training on athletic performance, we relied on A-V profiles as provided by Morin, Le Mat, Osgnach, et al. 2021 but in a slightly different way. Individual A-V profiles were modelled for each game. From the raw velocity $\vec{V}$ and a sampling frequency ω, we derived an acceleration vector $\vec{A}$ such that

$$A_i(nT) = V_i(nT) - V_{i-1}(nT)\,, \quad T = 1/\omega \quad \text{and} \quad \omega = 10Hz\,.$$

Here, we consider a signal $x(t) \rightarrow x\,[n]$ with $x\,[n] = x(nT)$ being the discrete formulation of $x(t)$.

Then, a first-order Butterworth filter was applied to the acceleration signal with a cut-off frequency of 1 Hz. Velocity observations were binned into $0.1\,\mathrm{m.s^{-1}}$ width bins in which the maximal acceleration values were retained. Hence, we modelled A-V profiles over velocities superior to $3\,\mathrm{m.s^{-1}}$ using a linear regression between acceleration and velocity (see Figure 3.9). A total of 1032 profiles were modelled, for an average equal to 25.80 ± 20.37 per player. The large standard deviation is related to some occasional players who only played a few games through the season.

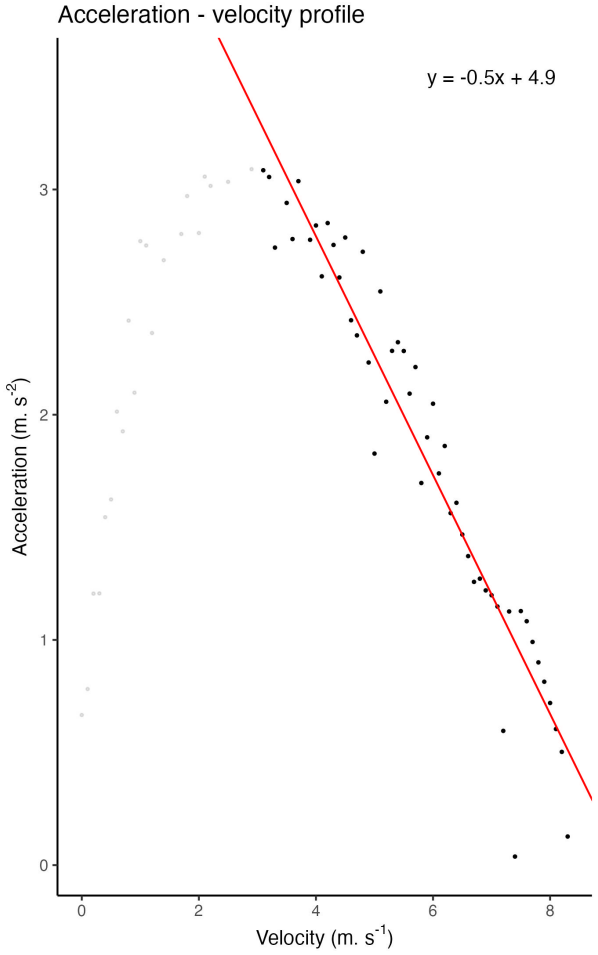
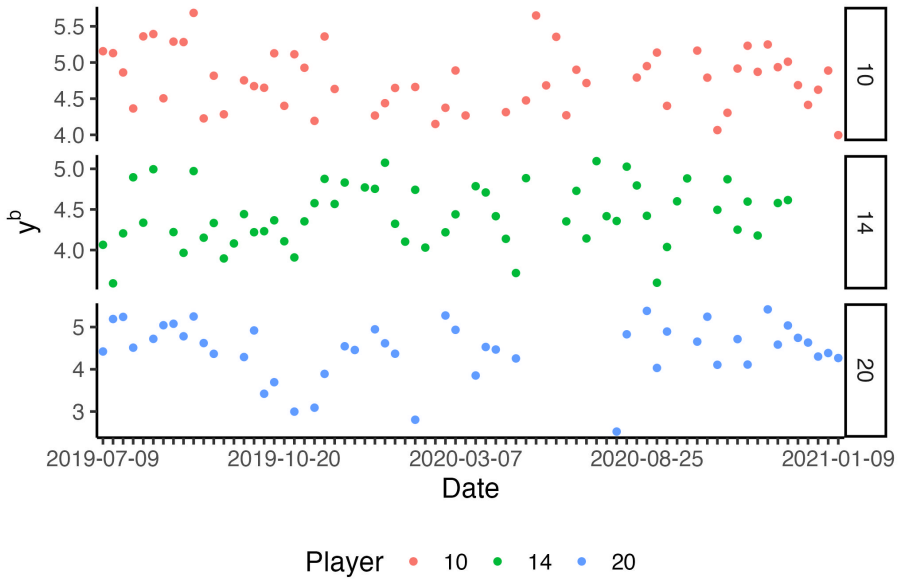
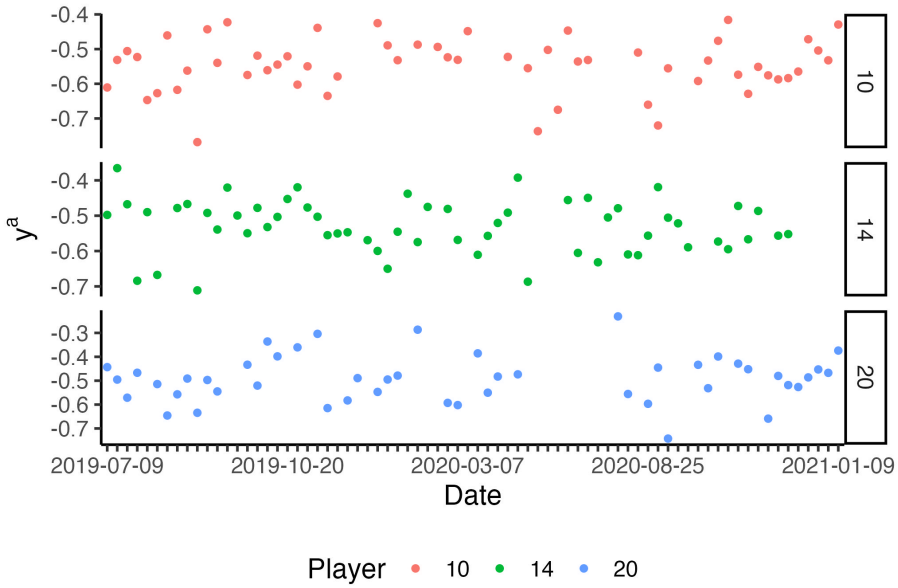


Figure 3.9. – Example of A-V profile modelled for a given player and a randomly selected game. Only plain dots were used for fitting the linear regression.

The performance definition is given by , $\left\{Y^a, Y^b\right\} \in \mathcal{Y}$ such that Y^a and Y^b refer to the corresponding slope and the intercept of individual A-V profiles respectively. Therefore, each observation in the ensemble $\left\{y^a_{j,t}, y^b_{j,t}\right\} \in \left\{Y^a, Y^b\right\}$ is related to both an athlete i and the day of realisation t. A sample of fitted coefficients is presented in Figure 3.10.



(a)



(b)

Figure 3.10. – Evolution of A-V profiles fitted intercept and slopes over the period of study. Players are randomly selected.

To formalise, allowing that $X \times Y \sim f$ with a density function f, the built data set is a sample $S = \left\{ \left(x_j, y_j^p\right) \right\}_{j \leq n} \sim f^n$.

3.4.2. Predicting the acceleration-velocity profile from GPS summarised features

Changes in A-V fitted parameters were investigated using two distinct prediction approaches: regression and classification. For both of them, linear models (a ridge regularisation for regression task, and a linear support vector classification (lSVC) for the classification one) used features pooled according to the two aforementioned aggregation methods (see Equations 3.4 and 3.5) and were compared to a LSTM, a particular case of RNN. For the same reasons evoked in Section 3.2.2.2, ridge regularisation allows solving the collinearity problems in multivariate modelling and therefore to prevent biased estimates through penalising estimates of correlated features (Hoerl and Kennard 1970; Marquardt and Snee 1975). According to the two aggregation methods, the multivariate linear models $m^{ridge} : X_p \rightarrow Y$ and $m^{ridge^*} : X_{p^*} \rightarrow Y$ take the general formulation

$$y_t^{(\text{ridge})} = \mathbf{x_t}^t \beta + \epsilon_t \ ,$$

Where $\mathbf{x} \in X_p$ the predictors pooled according to the *mean* function (see Equation 3.4) and $\mathbf{x} \in X_{p^*}$ the predictors pooled according to the exponential weighting function (see Equation 3.5) for m^{ridge} and m^{ridge^*}, respectively. Also, $\beta \in \mathbb{R}^d$ denotes the parameters of the model and ϵ_t the random error term.

RNN is the class of neural networks that consider past information to be used as inputs while having hidden states. Let consider a multidimensional vector $\mathbf{X}$ of fixed length l and dimension d, which includes unpooled summarised features as model input. Basically and from a $l \times d$ matrix, a RNN layer successively combines the current values of $\mathbf{X}_t$ of size d with the predicted value at time $t-1$ to return an output $\mathbf{h}_t$, defined by a function $f(\mathbf{X}_t, \mathbf{h}_{t-1})$ (see Figure 3.11a). This procedure is repeated as many times as there are RNN layers. However, RNN suffers from a short-term memory due to a vanishing gradient problem. Yet used for updating neural network weights, a gradient that shrinks as it back propagates through time stops the learning of layers. These layers may thus cause a loss of past information, particularly the long term one.

Introduced by Hochreiter and Schmidhuber 1997, LSTM aims at conserving the long term information through extended internal mechanisms. The LSTM architecture benefits from a cell state and various gates that regulate the flow of information. As shown in Figure 3.11b the cell state maps the previous cell state $\mathbf{C}_{t-1}$ to a new cell state $\mathbf{C}_t$ in which all the relevant information is carried throughout the sequence and where gates add or removes information from it. We can summarise the LSTM procedure such as:

1. A forget gate with a sigmoid activation function products a vector $\mathbf{f}_t$ of size $\mathbf{R}$ that contains relevant past information and values between 0 and 1. According to A (see Figure 3.11), a term-by-term multiplication between $\mathbf{f}_t$ and $\mathbf{C_{t-1}}$ act as a filter over the cell state with

$$\mathrm{A} = \mathbf{f}_t \otimes \mathbf{C}_{t-1} \,.$$

Here, $\otimes$ denotes the tensor product of two vector spaces.

2. Then, the old cell state $\mathbf{C}_{t-1}$ is updated according to a candidate vector

$$\mathrm{B} = \mathbf{i}_t \otimes \mathbf{k}_t \,.$$

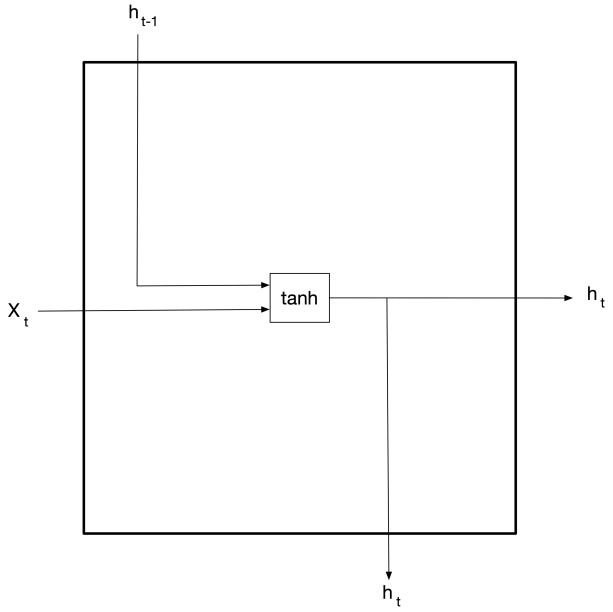
3. Accordingly, summing A and B let produce a new filtered cell state C_t, such as

$$\mathrm{C} = \mathbf{C}_t = (\mathbf{f}_t \otimes \mathbf{C}_{t-1}) + (\mathbf{i}_t \otimes \mathbf{k}_t) \,.$$

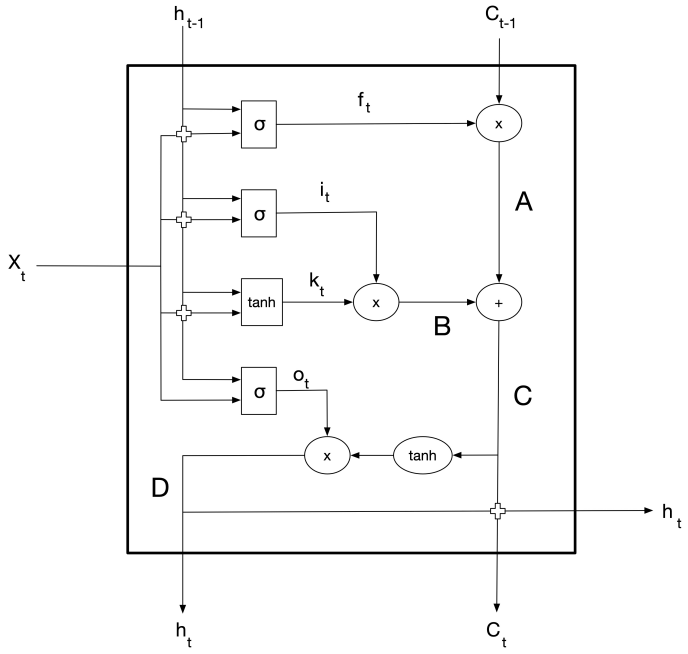
4. Analogously with $\mathbf{f}_t$ and $\mathbf{i}_t$, an output gate produces a filter $\mathbf{o}_t$ of values between 0 and 1 and size $\mathbf{R}$. Values of the new cell state $\mathbf{C}_t$ are returned to a $]-1, 1[$ interval according to a *tanh* activation function. The output $\mathbf{h}_t$ becomes

$$\mathrm{D} = \mathbf{h}_t = \mathbf{o}_t \otimes \tanh(\mathbf{c}_t) \,.$$

After all, the model is a function $m^{LSTM} : X \to Y$.



(a)



(b)

Figure 3.11. – Simplified diagram of (a) a RNN cell and (b) a LSTM cell.

Pre-processing.

Prior modelling, predictors were scaled according to a *Min-Max normalisation*. From a given feature X_d, we write the scaled X'_d such that

$$X'_d = \frac{X_d - min(X_d)}{max(X_d) - min(X_d}) \,. \tag{3.7}$$

3.4.2.1. Time-series forecasting

Generalities.

In Section 3.2.3, we showed that the auto-regressive component of the target variable was influential in the prediction of individual athletic performances. Hence, we started the modelling by defining baseline prediction performances given by time-series forecasting, using observations from games only.

In time-series forecasting, models without covariates use a restricted data set in which predictors are merely time information. The forecasting is made from information found in trend and seasonality components. In order to found the most performing models for time-series forecasting, we proceeded with a model selection using a simple holdout procedure, according to a split ratio of 0.8 (see 3.1.2.1 for details). One can righteously expect a linear relationship between changes in theoretical maximal acceleration and maximal running velocity. Hence, the ensemble $\{Y_a, Y_b\}$ was predicted in two different ways: sequentially (uni-modal forecasting) and concomitantly (multi-modal forecasting).

Afterwards, we benefited from the selected forecasting models by combining them into a weighted average ensemble for better performances than a randomly selected single model on average.

Forecasting results.

As shown in Figure 3.12, we observed different MAPE values between intercept and slopes predictions of A-V profiles. A two-ways repeated measure ANOVA confirmed these differences with a significant lower MAPE for the predictions given by profiles intercept than ones by profiles slope ($\beta_{diff} = -0.036 \in [-0.053, -0.020]\, 95\,\% CI, p < 0.001$, partial $\eta^2 = 0.28 \in [0.09, 0.46]\, 95\,\% CI$). After having linearly re-scaled MAPE values due to different range and variance between intercept and slopes (averaged range = 0.325 and range = 3.98; $\sigma^2 = 0.006$ and $\sigma^2 = 0.246$ for the slope and the intercept, respectively), a slight trend for an easier prediction task on A-V intercept was observed ($\beta_{diff} = -0.011 \in [-0.07, 0.003]\, 95\,\% CI, p = 0.122$).

Average ensembles were built following a model selection of a large set of time-series forecasting models $\mathcal{M}_{ts}$. In the uni-modal approach, the forecasting models which provided the lowest MAPE in prediction were Prophet (Taylor and Letham 2018), Theta and four Theta (Assimakopoulos and Nikolopoulos 2000), and fast Fourier transform (FFT) based. As expected, the combination of these models into an averaged ensemble provided the best performances (see Figure 3.13 for examples). In the multi-modal approach, the retained forecasting models were vector autoregressive moving average (VARIMA) (Tiao and Box 1981), RNN-LSTM,

and auto-regressive encoder-decoder Transformer (Vaswani, Shazeer, Parmar, et al. 2017). In this case, RNN-LSTM as well as averaged ensemble provided the best performances for predicting A-V slopes and intercepts, respectively. However, averaged ensemble provided only a slight trend for a greater accuracy and were not significantly more accurate than single forecasting models on average ($\beta_{diff} = -0.002 \in [-0. - 0.020, 0.012]\, 95\,\% CI, p = 0.86$).

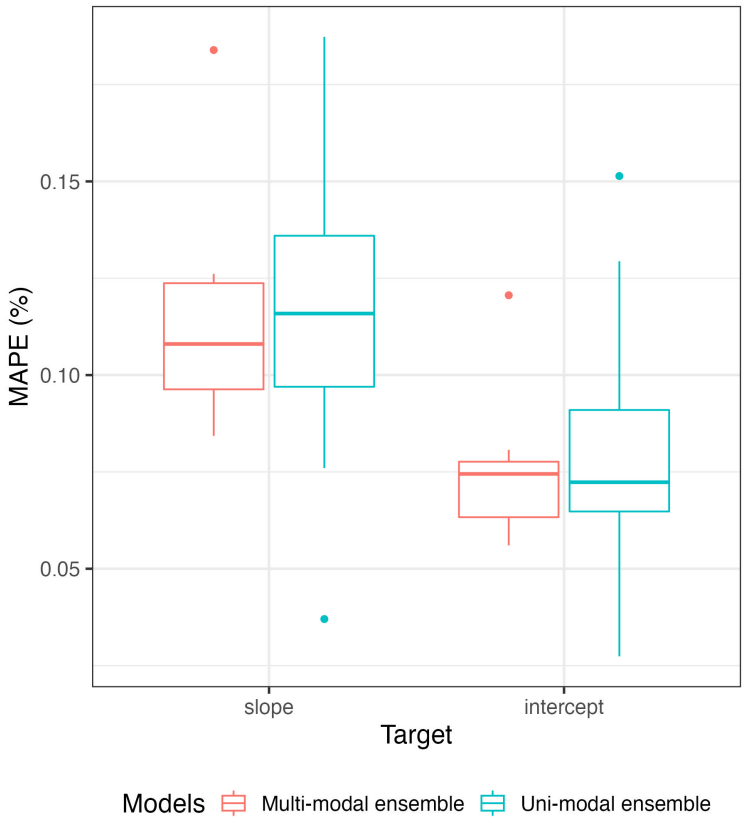
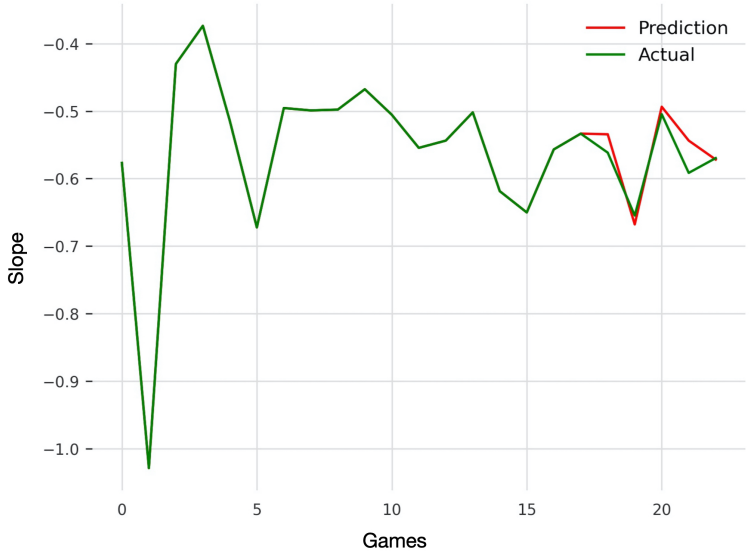
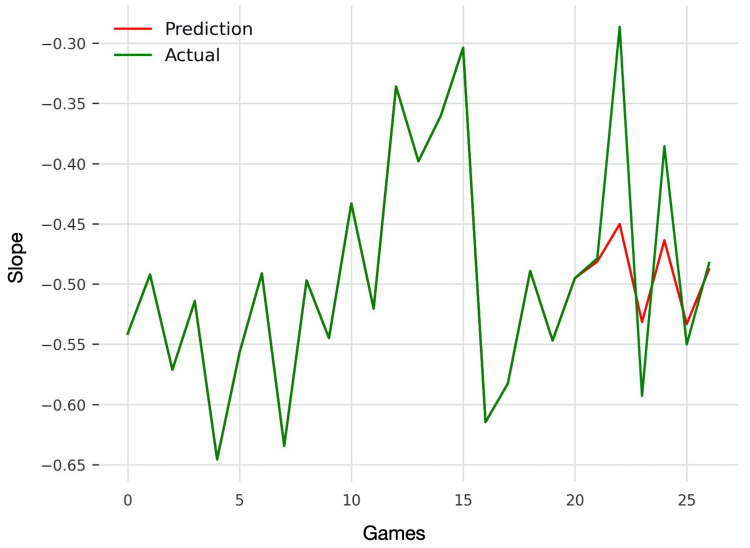


Figure 3.12. – Distributions of averaged time-series forecasting performances.

Performing multi-modal forecasting might provide better results, but it also requires a larger sample size than uni-modal forecasting methods to estimate model parameters correctly. Accordingly, we filtered out players who performed less than forty games for computing multi-modal forecasting models. Only nine players were retained for the multi-modal forecasting task, whereas the uni-modal task included data from nineteen players. Therefore, the sample size heterogeneity should be considered when interpreting the forecasting results presented so far since a larger sample size might reasonably yield better or at least different forecasting performances. Overall synthesis of the selected forecasting models and their performances is displayed in Table 3.5.



(a)



(b)

Figure 3.13. – Example A-V profiles slopes forecasting using the uni-modal averaged ensemble. (a) represents the best prediction, (b) is the median prediction.

Table 3.5. – Average MAPE for each selected model. (*) additive seasonality, (**) multi-modal models required longer time-series. We limit the study of these models to time-series larger than 38 observations.

Models	MAPE_{slope}	$\text{MAPE}_{intercept}$	Multi-modality**
Prophet	0.134	0.095	✗
Theta	0.150*	0.096	✗
FourTheta	0.120	0.085	✗
FFT	0.161	0.121	✗
Ensemble	**0.115**	**0.081**	✗
VARIMA	0.162	0.127	✓
RNN-LSTM	**0.111**	0.099	✓
Transformers	0.120	0.075	✓
Ensemble	0.113	**0.072**	✓

The forecasting methods presented in this section serve as a reference for further performance predictions and model comparisons. In practical terms, the main limitation of using such forecasting models is that we only consider data from games for making previsions. Hence, interpretations drawn from each forecast are restricted to the effect of preceding games on the next game, and the contribution of any training session preceding a performance remains hidden.

In order to investigate more precisely the effect of the training sessions on the changes in athletic performance, multivariate analysis including data from training sessions is required. In the following sections, we provide some applications using the data defined in Section 3.4.1.2 and displayed in Table 3.4.

3.4.2.2. Multivariate modelling using commercial features

As presented in Section 3.4.1, we aimed at predicting $\{Y_a, Y_b\}$ using two sets of aggregated predictors X_p and X_{p^*} from the original features displayed in Table 3.4. Since models rely on several predictors, we consider the modelling being multivariate. In addition, predictions were made according to regression and classification tasks.

Regression task.

Similar to the time-series forecasting and for any model, distributions of MAPE were significantly lower when predicting the A-V model intercept than the slope ($\beta_{diff} = -0.039 \in [-0.047, -0.031]\,95\,\%\,CI, p < 0.001$, partial $\eta^2 = 0.32 \in [0.22, 0.41]\,95\,\%\,CI$). The re-scaled MAPE also provided significant differences in favour of the intercept predictions ($\beta_{diff} = -0.013 \in [-0.021, -0.005]\,95\,\%\,CI, p = 0.002$, partial $\eta^2 = 0.06 \in [0.01, 0.15]\,95\,\%\,CI$). Post-hoc comparisons showed that there is a slight trend for a better performance when predicting the A-V intercept in favour of both LSTM and ridge regression ($\beta_{diff} = -0.013 \in$

$[-0.028, 0.002]\,95\,\%\,CI, p = 0.12$ and $\beta_{diff} = -0.013 \in [-0.029, -0.002]\,95\,\%\,CI, p = 0.103$, respectively). Besides, there is no significant difference in errors distribution among models for each target ($p = 0.165$). That suggests that using data from training sessions within multivariate modelling did not improve the prediction of A-V coefficients compared to time-series forecasting models (either uni-modal or multi-modal averaged ensemble models, see Figure 3.14)

Except for the latter, which only considered data from games (see Section 3.4.2.1), there was no advantage of using an exponentially weighted aggregation (refer to Equations 3.5 and 3.6 for details) over a simple aggregation according to the mean ($\beta_{diff} = -0.001 \in [-0.012, 0.010]\,95\,\%\,CI, p = 0.979$). Hence, we consider the accuracy of models to be equal regardless of the aggregation methods used. It let us point to either a limited relevance of the current explanatory features or weak predictability of the A-V profile using GPS features only. Note that we excluded irrelevant features from the modelling by employing a feature selection based on the linear correlation between the feature of interest and the target (*i.e.* according to the F value). Accordingly, we held the ten most meaningful features for making further predictions.

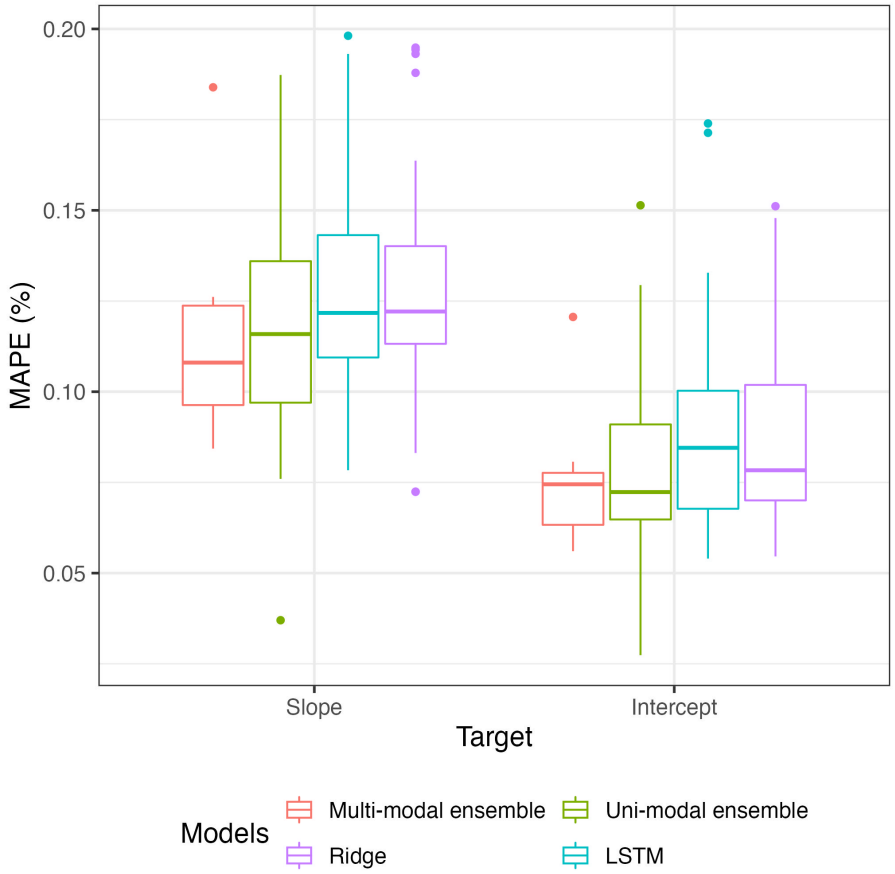


Figure 3.14. – Distributions of model performances in a regression task.

On average, individually fitted models did not provide lower prediction errors than those fitted on the group ($p = 0.468$). Therefore, this suggests that there is no benefit of using player-specific models for predicting A-V profiles with the current data.

Classification task

In parallel, the problem might also be formulated as a classification task. The coefficients of A-V profiles were binned into classes, according to a *uniform* clustering and a *kmeans* clustering (Hartigan and Wong 1979). *Uniform* clustering let us bin A-V profiles coefficients into horizontally scaled classes. This is a simple approach which considers fixed values (*i.e.* in our case the quartile distribution) for determining classes, independently of the probability density function of the variable.

Alternatively, a *kmeans* clustering algorithm determines classes according to k identified number of centroids (*i.e.* expected locations representing the center of the clusters). Then, each observation is allocated to a particular cluster for which

the RSS of the distance between the point and the centroid is minimised. Therefore, the *kmeans* clustering is more sensitive to the data than the *uniform* clustering.

For both methods, we empirically set the number of classes to $k = 4$. That was motivated by the fact that four classes are sufficient to build a relevant representation of muscular and athletic properties while helping in the prediction problem.

A summary of models performances for the classification task is provided in Table 3.6. In average, the $F1 = 2\frac{\text{precision.recall}}{\text{precision+recall}}$ (F1-score) of the models was not significantly different between intercept and slope predictions ($p > 0.05$). However, we found significant differences between models where lSVC was prone to a slight better F1-score than LSTM ($\beta_{diff} = 0.021 \in [0.001, 0.041]\,95\,\%\,CI, p = 0.041$). Whether lSVC or LSTM, group-computed models provided greater performances than ones fitted for individuals ($\beta_{diff} = 0.146 \in [0.101, 0.192]\,95\,\%\,CI, p < 0.001$, partial $\eta^2 = 0.02 \in [0.01, 0.04]\,95\,\%\,CI$). The *uniform* clustering provided better F1-score than *kmeans* clustering ($\beta_{diff} = 0.023 \in [0.003, 0.043]\,95\,\%\,CI, p = 0.023$). However, we note that the effect size reported for such differences is not significant (partial $\eta^2 = 0.01 \in [0.00, 0.01]\,95\,\%\,CI$), suggesting a slight trend only for better performances using a *uniform* clustering.

Finally, there is no differences were observed among aggregation methods of time-series predictors (*i.e. mean* or exponentially weighted aggregations, , $p > 0.05$).

Regarding the averaged model performances, a classification approach to the problem is better suitable using one model for all players. However, the classification task remains not advantageous compared to the regression.

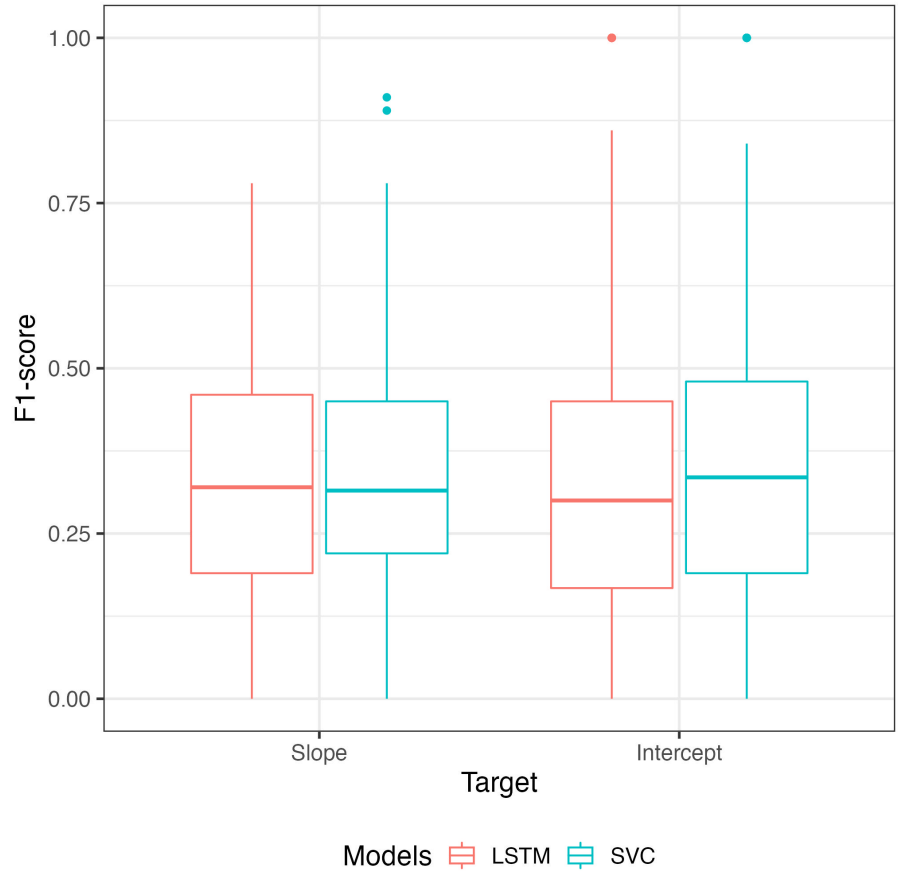


Figure 3.15. – Distributions of model performances in a classification task.

Table 3.6. – Summary of models performances for the classification task

Model	Population	Precision	Recall	F1-score
LSTM	global	0.434	0.557	0.463
LSTM	individual	0.319	0.383	0.316
SVC	global	0.446	0.563	0.482
SVC	individual	0.360	0.378	0.337

Conclusion

The results presented so far show that the use of commercial features through multivariate modelling does not outperform time-series forecasting models, which were strictly based on time and target information without any other covariates. Hence, it questions either the relevance of GPS commercial features for modelling the effect of training on A-V profiles or the predictability of such profiles using these features only.

On this basis, it could be worth investigating the extraction of *new* features directly from raw GPS data.

3.4.2.3. Extracting new features from raw global positioning system data

As described in Section 3.4.1.2, the player position was recorded by GPS at a sampling frequency of 10 Hz. Timestamp, player position (*i.e.* latitude, longitude) and velocity were stored in a comma separated values file for each player and session. Since commercial features were computed from the raw velocity vector $\vec{V}$ and its derivative $\vec{A}$, we processed a feature extraction directly from $\vec{V}$.

First, we consider $\vec{V}$ being a stationary time-series $(X_t)_{t\in\mathbb{R}}$. Formally, a time series is stationary if the law $\mathcal{L}$ of any vector $(X_{t_1}, \ldots, t_n)$ is time translation invariant, that is $\mathcal{L}(X_{t_1}, \ldots, t_n) = \mathcal{L}(X_{t_1+h}, \ldots, X_{t_n+h}), \quad \forall (t_1, \ldots, t_n) \in \mathbb{R}^n \quad and \quad h \in \mathbb{R}$ with t being a time value and $\mathbb{R}$ being a set of real numbers (Cox and Miller 2017). The stationary of time-series was checked using a Dickey-Fuller test (Fuller 2009).

Several features were extracted from the time-series in both time and frequency domains through FFT. For this purpose, we used the *tsfresh* Python module (Christ, Braun, Neuffer, et al. 2018). The feature extraction from both domains provided categorized 779 features Christ, Braun, and Neuffer n.d. A feature selection like performed during the previous tasks let us retain only the ten most relevant features, according to their significance level (F statistic and p value).

Distributions of raw prediction errors was provided in Figure 3.16. There is no significant difference reported between averaged intercept and slope predictions. Only a slight trend for a lower MAPE was imputed to intercept predictions ($\beta_{diff} = -0.012 \in [-0.027, 0.002]\, 95\,\% CI, p = 0.09$). In addition, individual and group computed LSTM provided similar performances in terms of accuracy ($p > 0.05$). That indicates that there is no substantial advantage of computing models per player.

Based on the previous results, a multilevel comparison of predictions across models (see Figure 3.17) did not show any significant difference between uni-modal and multi-modal time-series forecasting models, multivariate models using commercial features and multivariate models using features extracted from raw data.

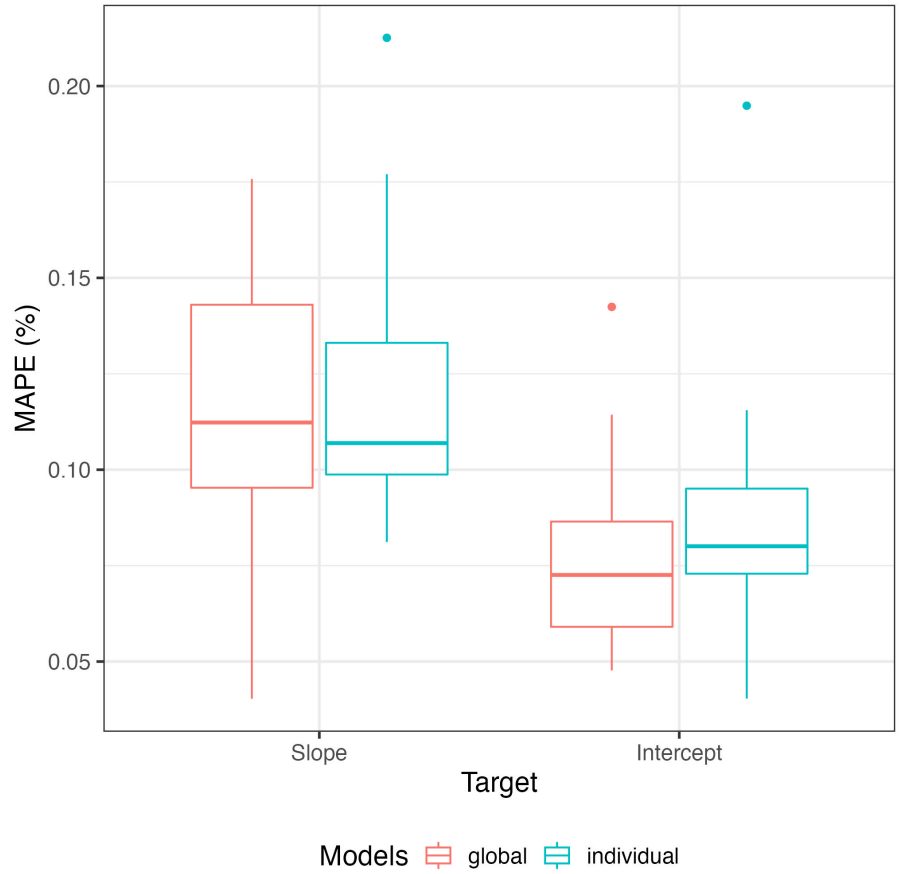


Figure 3.16. – Distributions of model performances in a time-series forecasting using features extracted from raw data.

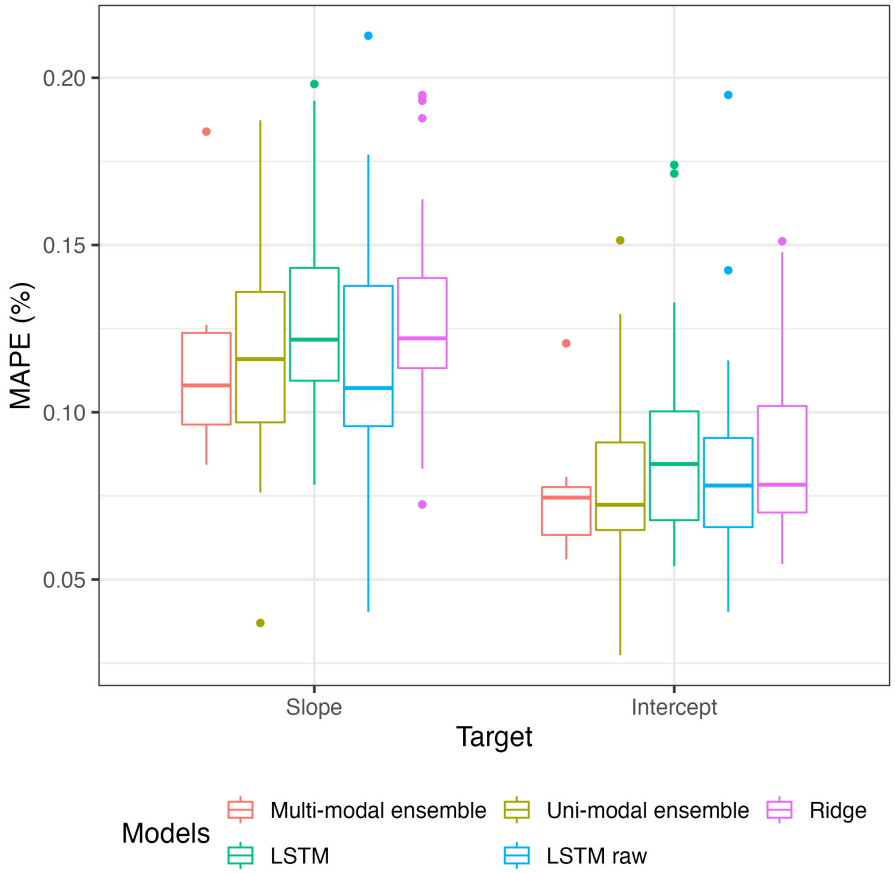


Figure 3.17. – Distributions of model performances across models.

Table 3.7. – Summary of models performances according to intercept and slope coefficients. $\overline{MAPE}$ represents the averaged MAPE over individuals and validation folders. The population represents either models computed over the group of players (G) or individually computed models (I).

Model	Target	Population	Aggregation	$\overline{MAPE}$
Multi-modal ensemble	intercept	I	N/A	**0.076**
LSTM (raw)	intercept	G	N/A	0.077
Ridge	intercept	G	exponential	0.080
Ridge	intercept	G	mean	0.080
Uni-modal Ensemble	intercept	I	N/A	0.080
LSTM	intercept	G	exponential	0.084
LSTM	intercept	G	mean	0.084
Ridge	intercept	I	mean	0.085
Ridge	intercept	I	exponential	0.085
LSTM (raw)	intercept	I	N/A	0.088
LSTM	intercept	I	mean	0.089
LSTM	intercept	I	exponential	0.090
LSTM	slope	G	mean	**0.114**
LSTM	slope	G	exponential	0.114
Uni-modal Ensemble	slope	I	N/A	0.115
Multi-modal ensemble	slope	I	N/A	0.116
RIDGE	slope	G	mean	0.116
RIDGE	slope	G	exponential	0.116
LSTM (raw)	slope	G	N/A	0.119
LSTM (raw)	slope	I	N/A	0.121
LSTM	slope	I	mean	0.126
RIDGE	slope	I	mean	0.128
LSTM	slope	I	exponential	0.128
RIDGE	slope	I	exponential	0.129

3.4.3. Conclusion

In this study, we aimed at modelling coefficients of individual A-V profiles. For this purpose, we first considered time-series forecasting models, which used data from games only as of the baseline of models performances. Then, multivariate modelling approaches were compared to these baseline models with a regression task - using regularised linear regression (ridge) and LSTM - and a classification task - using lSVC and LSTM. For the linear regression and lSVC, predictors were aggregated according to two distinct functions: a *mean* and an exponential weighted function (both of them are defined in Equations 3.4 and 3.5). Finally, we extracted new features from GPS raw data and assessed their contribution to the modelling.

We recall that excepted for time-series forecasting models, models were fitted either per player or over the group of players.

As provided in Table 3.7, they were only slight differences in terms of accuracy among models. Using commercial features did not outperform time-series forecasting models regardless of the aggregation method used. It first questions the relevancy of the commercial features provided by the sensor that is built from sport science experts and related to some physiological adaptations to training (Hader, Rumpf, Hertzog, et al. 2019). It also questions the predictability of such A-V profiles recently proposed by authors (Morin, Le Mat, Osgnach, et al. 2021).

Besides, features extracted from raw data in both time and frequency domains provided similar results using the commercial features. Even though not being based on expert hypotheses but fully extracted from signal processing methods, they appeared to be as valuable as the commercial ones and opened the way to future works on data mining and knowledge discovery in the sports field. One issue related to this perspective could be the direct interpretability of the extracted features, particularly those related to the frequency domain. However, the growing technology provides high sampling frequency systems (*e.g.* IMU and motion capture), from which a physiological representation of response to exercise might be extracted. Also, going through raw data of IMU might contribute to solving the enigma of injury occurrence, which remain a hot research area in sports science.

According to intercept and slope coefficients and considering re-scaled MAPE, the accuracy of model predictions varied between 7% and 10%. We believe that this is an acceptable accuracy since the A-V profile depends on un-measured and un-controlled factors, namely the opponent activity, then any psychological, environmental, nutritional aspects. In addition, the detection of interpretable key features might allow guiding field training and resistance training, following individual needs (*e.g* either maximal force or velocity enhancement). Nonetheless, further studies that account for broader sources of data would be recommended for validating the use of individual A-V profiles as a valuable performance criterion for athletic performance and injury prediction purposes. Since the maximal power is directly related to any sprint performance (Chelly and Denis 2001), alternatives considering force-power profiles and maximal power could be worth investigating.

3.4.4. Take home message

— Time-series forecasting models provided the greatest performances despite using data from games only.
— Given the multi-factorial aspect of A-V profiles, the use of only GPS data provided an acceptable MAPE. However, further investigations with broader sources of data would be required in order to validate the sensibility and relevancy of A-V as an athletic performance criterion.
— To use new features extracted from GPS raw data according to signal processing methods provided similar model performances compared with models

which accounted for commercial features. Even though the standard features are based on expert knowledge, raw data should be considered for future athletic performance and injury occurrence analysis.

4. A decision support system for team and athletes

Table of contents

4.1 Introduction . . . 190
4.2 Validation of a running power meter . . . 192
4.2.1 Methodological approach . . . 192
4.2.1.1 Population studied . . . 192
4.2.1.2 Protocol . . . 193
4.2.1.3 Calculations . . . 195
4.2.1.4 Ground Contact Time, Stride Time and Stride Frequency . . . 195
4.2.1.5 Time Matching . . . 197
4.2.2 Verification of reference measures . . . 197
4.2.2.1 Mechanical Cost of Running . . . 197
4.2.2.2 Ground Contact Time and Leg Stiffness . . . 198
4.2.3 Stryd and reference measures comparisons . . . 200
4.2.3.1 Consumed Metabolic Energy and Power Output . . . 200
4.2.3.2 Mechanical Power . . . 202
4.2.3.3 Ground Contact Time . . . 203
4.2.3.4 Leg Spring Stiffness . . . 207
4.2.4 Conclusion . . . 207
4.2.5 Take-home message . . . 208
4.3 *RunningPerfApp*: A *shiny* web-application for long distance runners 208
4.3.1 Monitoring athlete status . . . 208
4.4 Team management and data visualisation with SAP analytics cloud 211
4.4.1 Monitoring team and players body mass . . . 212
4.4.2 Monitoring player performances . . . 214
4.4.3 Monitoring training from global positioning system data . . . 215
4.4.4 Monitoring Well-being . . . 215

4.1. Introduction

To summarise, we first presented state of the art about knowledge in modelling training effects on athletic performance. Then, we focused on TL quantification

methods in RT in order to provide valuable and meaningful features based on physiological observations. These features are intended to be monitored and considered as model input when dealing with RT in amateurs and athletes. The previous chapters came with two predictive modelling applications, one providing a robust methodology according to model generalisation in individual elite sports. The other showed how athletic performance could be modelled from GPS measurements. Throughout all these applications, we emphasise the multidimensional and complex nature of athletic performance, for which a systemic modelling approach is highly recommended.

The massive cost of injuries in sports has motivated sports teams, instances and private companies to engage in financial endeavours in research and development of injury prevention tools. As part of a company (*Seenovate*, France), we provided in this chapter some DSS developments intended for amateurs and professional teams. Objectives behind DSS are to support customers in their decision-making about athletic performance and injury problems. To do so, we rely on a general and non-exhaustive framework such as presented in Figure 4.1.

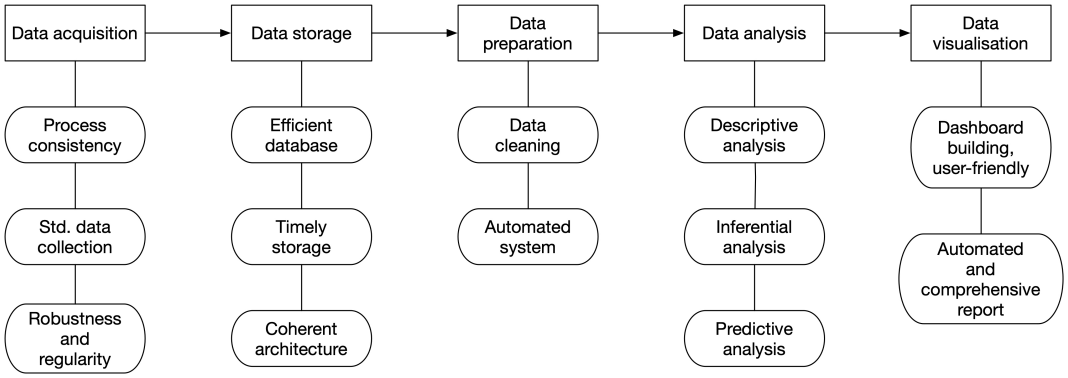


Figure 4.1. – Basic framework for DSS building in sports.

The first and crucial step of any DSS development starts with the acquisition of data. In order to ensure that data are correctly collected and gathered, a few key points might be underlined. First, data should be measured within consistent and standardised processes to rule out any possible methodological biases. Any data omission might be highly detrimental to further data analysis. Then, convenient data storage comes while organising data into an efficient database, accessible and with an architecture designed for further modelling. Preparing data for analysis is a large part of the problem. It consists of several data pre-processing, which include data verification, signal processing, and data transformation. This step should be as far as possible automated in order to standardise the process and to deal efficiently with new data entries. Data analysis are thereupon possible and may be identified into:

1. Descriptive analysis in which we aim at discriminating data and let emerge relationships and hypotheses.

2. Inferential analysis which give information about features of interest and allows for drawing interpretations from observed relationships.
3. Predictive analysis for predicting an athletic performance at a particular time, processing simulations of training load sequences and extracting key insights intended for coaches, trainers and medical staff.

The last point is directly related to the data visualisation step. It intends to make the models output, predictions, statistical inferences comprehensible for all through the development of user-friendly interfaces and dashboards.

In summary, data is at the basis of any DSS. Even if they are consistently collected, users should be assured that the measurements they use are valid and reliable. As an initial preoccupation and before DSS development, we assessed the validity of a promising wearable sensor applied to runners.

4.2. Validation of a running power meter

The Stryd power meter (Boulder, Colorado, United States) is a pioneer in the field. It provides the following measures in real time: pace, running PO, vertical oscillation, elevation, distance, GCT and leg spring stiffness (LSS). Since the algorithms used to derive the metrics from the Stryd sensor are proprietary and despite some investigations of its validity (Garcı+a-Pinillos, Latorre-Román, Roche-Seruendo, et al. 2019; Garcı+a-Pinillos, Roche-Seruendo, Marcén-Cinca, et al. 2018; Lara, Shearer, Coppi, et al. 2018; Navalta, Montes, Bodell, et al. 2018), the estimate of these metrics remains unclear. To our knowledge, these aforementioned studies compared neither PO estimates nor LSS estimates to reference systems. Thus, their absolute and relative validity remains to be explored. Therefore, PO, LSS and GCT measures from the Stryd foot pod at different sub-maximal running speeds were compared to reference systems in ecological conditions. To do so, validated and reference methods were used to calculate the $\dot{W}_{ext}$ and the leg stiffness from accurate force platform measures (Cavagna 1975; McMahon and Cheng 1990).

4.2.1. Methodological approach

To assess the validity of the Stryd power meter, we compared the recorded data with force platforms, a motion analysis system and a portable metabolic analyser on a 200 m indoor running track.

4.2.1.1. Population studied

Six recreational runners (four male age: 39 ± 3 years, $\dot{V}O_2max$: $53.85 \pm 6.09\,mlO_2 \cdot min^{-1} \cdot kg^{-1}$ and two females age: 35, $\dot{V}O_2max$: $48.33 \pm 6.75\,mlO_2 \cdot min^{-1} \cdot kg^{-1}$) voluntary participated in this study. All participants validated the inclusion criteria:

(i) older than 18 years; (ii) train 3 to 5 times a week and on a treadmill at least once per week; (iii) not suffering from any injury impacting running capacities for the last 6 months.

4.2.1.2. Protocol

The test consisted of an incremental running trial around a 200m track (Figure 4.2). The initial speed was set at 10 km/h and 8 km/h for men and women, respectively. After that, the speed was increased by $0.5km.h^{-1}$ every minute. Cones were set at 20m intervals along the 200m track (inside the first line). An audio signal dictated the running pace and the runners had to be within 2m of the cones at each beep signal. When a runner was behind a cone three consecutive times, the test stopped. Individual MAS was determined as the lowest speed at which $\dot{V}O_2max$ was attained (Billat, Hill, Pinoteau, et al. 1996). MAS was reached between 12 and 20min of exercise in order to limit impairments caused by the accumulation of fatigue (allowing MAS values to reach approximately $20km.h^{-1}$ at 20min). The participants were rested before the start, the first minutes of the test acting as a warm-up.

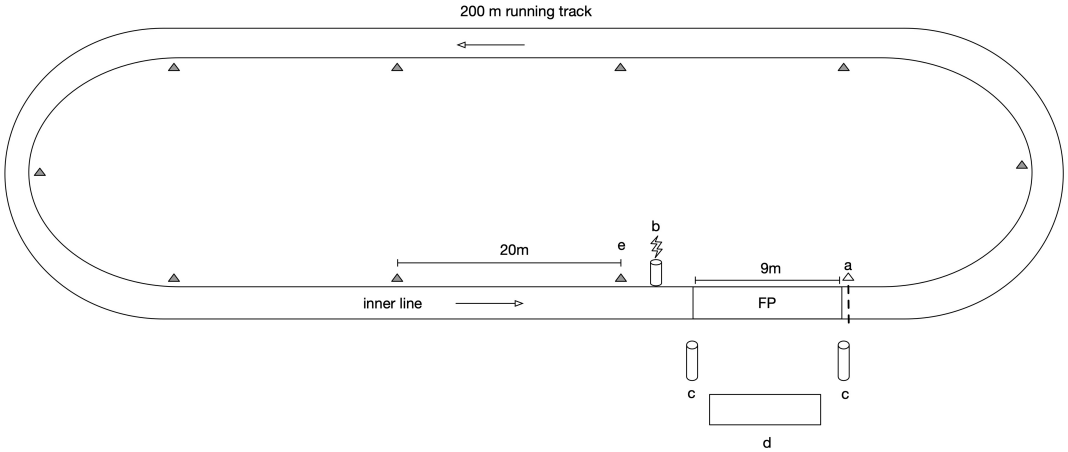


Figure 4.2. – MAS test protocol on a 200m indoor track. The symbol a represents the start line, b is a photoelectric cell to reset the force platform records, both c are the two motion analysis sensor modules, d is the control panel and e are the cones laid every 20m and FP is the force platform recording area.

Power meter

Each participant wore the Stryd power meter, a foot-mounted inertial sensor of 9.1 g reinforced with carbon fibre, firmly attached to the shoe according to the manufacturer's recommendations. The device stores at 1 Hz sampling frequency the following variables: GCT, vertical oscillation, running PO, distance, LSS, cadence. According to Stryd team information, the device is operational out

of the box and should not need any calibration, accepting a measurement error of 3%. Participants filled in their body height and body mass prior to its use, requisites for the PO estimation. As a precautionary measure, the device was fully loaded and activated 20min before the beginning of the test. The firmware version used was 1.1.9 (released in February 2019), and data were extracted in flexible and inter-operable data transfer (FIT) format from the Stryd application (`http//www.stryd.com/powercenter`).

Gas exchange measures

Participants wore the Cosmed k4b2 portable metabolic system in order to record breath-by-breath gas exchange measures. The device has been validated by several independent authors (Brisswalter and Tartaruga 2014; Doyon, Perrey, Abe, et al. 2001; Duffield, Dawson, Pinnington, et al. 2004). The device was checked and certified valid by the Cosmed company two months before the study. Before each testing session, the metabolic analyser was powered on to warm up for 30 min, and then CO_2 and O_2 sensors were calibrated based on known gas tank concentrations and ambient air measurements. Flowmeter calibration was then completed using a 3.0 L syringe according to the manufacturer recommendation. Calibration was done for each subject after 30 min warm-up activation. To produce uniform sampling for subsequent numerical analysis, we linearly interpolated data on a second-by-second basis. Due to a noisy signal provided by the portable metabolic system, data were averaged with a fifth-order moving average filter (Smith 2013), corresponding to 5 s time bins (Keir, Murias, Paterson, et al. 2014).

Force platforms

A track embedded force platforms system measured the ground reaction forces (GRFs) once per lap (200m) during the incremental running test. The system consisted of three force platforms (one Kistler, Switzerland and two AMTI, USA) of dimensions 90 cm * 90 cm connected in series and covered with a tartan mat. Each platform was calibrated before the study. The sampling frequency was set at 500Hz.

Motion analysis

The entire runner stride measurements along the platform section were performed using the Coda Motion 3-D movement analysis system (Charnwood Dynamics Ltd., United Kingdom). The system was composed of marker devices, sensor modules and data analysis software. The marker devices consist of infrared emission markers. The sensor module is made up of three optical sensors, which capture 3-D position and orientation by tracking the markers in real-time. Two sensor modules were placed on both sides of the platform area, and two markers (CXS models) were firmly fixed to the heel and on the fifth metatarsal bone of the runner. The system delivers reliable real-time 3D measurements on the computer

screen with a 400 Hz sample frequency throughout the experiment. The data were processed with the Coda Motion Odin software platform.

4.2.1.3. Calculations

4.2.1.4. Ground Contact Time, Stride Time and Stride Frequency

Changes in vertical GRFs signal were used to detect foot strike and to estimate GCT, defined by the duration of the GRFs variation. GCT was defined as the duration of the vertical GRFs signal. Flight time was estimated from the Z axis heel marker changes monitored by the Coda Motion analysis system. Assuming that participants have equal stride properties between lower limbs, the stride time was approximated as following:

$$Ts = \frac{C_2 - C_1}{2},$$

with Ts the stride time in seconds, C_2 and C_1 respectively the time instances (s) of the second and first heel strikes recorded by the optical sensor located on the foot. Hence,

$$\omega = \frac{1}{Ts}$$

is the stride frequency (Hz).

External Mechanical Power, Mechanical Cost of Running and Mechanical Efficiency

We calculated the external power and the mechanical cost of running following an external energy summation approach. GRFs signals were computed in the anterior-posterior (x) and vertical (z) axis. We omitted the lateral axis due to its negligible contribution when running on a track (Cavagna 1975). First, the acceleration $\vec{A}$ (m/s^2) is decomposed on both z and x axes, respectively defined as

$$Az = \frac{(Fz - mg)}{m} \quad \text{and}$$
$$Ax = \frac{Fx}{m},$$

where Fz and Fx are the components of the force recorded (N), m is the mass of the subject (kg) and $g = 9.80665$ is the acceleration due to the gravity (m/s^2). Consequently, the speed at time i denoted $\vec{S_i}$ (m/s) is

$$\vec{S}_i = \int_{t=i-1}^{i} \vec{A}_t \, \mathrm{d}t + \vec{S}_{i-1} \quad \text{and}$$
$$\vec{D}_i = \int_{t=i-1}^{i} \vec{S}_t \, \mathrm{d}t + \vec{D}_{i-1},$$

with $\vec{D}_i$ the distance (m) at time i and t the time measurement according to the sampling frequency (2 ms). Potential, kinetic and total work (W_p, W_k and W_t) were calculated as

$$W_p = mg\,(\max D_z - \min D_z)$$
$$W_k = \frac{1}{2} m \left(\max S_x^2 - \min S_x^2\right)$$
$$W_t = W_p + W_k,$$

where z and x are the vertical and anterior-posterior axes, respectively. Finally, $\dot{W}_{ext}$ was calculated by

$$\dot{W}_{ext} = W_t \, \omega.$$

The mechanical cost of running ($\mathrm{J{\cdot}kg^{-1}{\cdot}m^{-1}}$) ($Cm$) was thus defined as

$$Cm = \dot{W}_{ext} \, S_x^{-1} \, m^{-1}. \tag{4.1}$$

Metabolic power (W) ($\dot{W}_{met}$, W) and net mechanical efficiency (%) (ME) were estimated from $\dot{\mathrm{V}}\mathrm{O}_2$ measurements using energy equivalent of O_2 as following

$$\dot{W}_{met} = \frac{\dot{V}O_2 \, m}{60} \, k$$

$$ME = \frac{\dot{W}_{ext}}{\Delta \dot{W}_{met}} \, 100, \tag{4.2}$$

with $\dot{\mathrm{V}}\mathrm{O}_2$ in $\mathrm{mlO_2{\cdot}kg^{-1}{\cdot}min^{-1}}$, k the energy equivalent for the consumption of $1\,\mathrm{mlO_2}$ for a value of 21.1 J (Di Prampero, Atchou, Brückner, et al. 1986; Livesey and Elia 1988) and $\dot{W}_{met}$, W the variation of $\dot{\mathrm{V}}\mathrm{O}_2$ above resting.

Leg Stiffness

Stryd LSS metric was compared with the reference method (McMahon and Cheng 1990) for assessing k_{leg}, using force platform measures and defined as

$$k_{leg} = \hat{F}_z \, \Delta L^{-1}, \tag{4.3}$$

where $\hat{F}_z$ denotes the maximal values of F_z and $\Delta L = \Delta y + L(1 - \cos\theta)$. In the latter, $\theta = \sin(vT_c/2L)$, Δy is the vertical displacement of the center of mass,

v is the forward velocity, T_c is the contact time and L is the initial leg length standardized at 0.53 of the subject's height.

4.2.1.5. Time Matching

The following procedure was used to match power meter and force platform measurements. After checking the length of the track (200 m following the inner line), participants started to run right after the force platforms set up, spreading out over 9 m (Figure 4.2). The recorded sequences were matched by subtracting 9 m to every 200 m estimated by the power meter, assuming that the Stryd device measures distance reliably (Navalta, Montes, Bodell, et al. 2019). Values were averaged over this distance and over the three force platforms to keep one value for each metric of interest.

Statistical analysis was carried-out in a bayesian framework. Details about the null hypothesis significant testing alternative and formulation of the model used are provided in Appendix H.

4.2.2. Verification of reference measures

4.2.2.1. Mechanical Cost of Running

Using the force platforms, we calculated the mechanical cost of running following Equation 4.1. Mean values of Cm were 2.36 ± 0.46 J·kg^{-1}·m^{-1}. By comparing the Cm with the increase of speed, we observed a moderate negative linear relationship for all participants ($R^2 = 0.66$, $[0.60, 0.71]$ 95% CI). Both speed (as a component of Cm calculation) and subject have shown an effect on the Cm measure with an extreme evidence ($BF_{10} > 100$). In terms of explained variance by participant effect, intraclass correlation coefficient (ICC) reported a moderate correlation ($ICC = 0.65$, $[0.34, 0.94]$ 95% CI)) supporting the individual differences in Cm. Highest values of Cm were found at low speeds (up to $3m/s$).

Mechanical cost values reported were in agreement with expected values when the total work is calculated by assuming no energy transfer between potential and kinetic energies (Candau, Belli, Millet, et al. 1998; Harris, Debeliso, and Adams 2003; Ito, Komi, Sjödin, et al. 1983; Willems, Cavagna, and Heglund 1995). High values of Cm at low speeds reveal inefficient running patterns, as described previously (Carrard, Fontana, and Malatesta 2018; Harris, Debeliso, and Adams 2003; Ito, Komi, Sjödin, et al. 1983).

Metabolic and External Mechanical Power Relationship

$\dot{V}O_2$ and $\dot{W}_{ext}$ revealed a strong and positive linear relationship for the 6 participants ($R^2 = 0.85$, $[0.76, 0.89]$95%CI). Results of the LMM (Table 4.1) showed a significant effect of mechanical power over $\dot{V}O_2$ ($\beta_1 = 0.081$, $[0.035, 0.111]$95%CI).

Bayes factor supported the results with an extreme evidence of both $\dot{W}_{ext}$ and subject effect ($BF_{10} > 100$). Large standard deviations of intercept parameter indicated heterogeneous levels between participants. ICC supported the individual differences with a high correlation ($ICC = 0.94$, $[0.55, 1]$ 95% CI)). The net ME was calculated from the force platform measures following Equation (4.2). Group mean value of ME was equal to $55 \pm 3\%$.

Table 4.1. – Linear mixed modelling of $\dot{V}O_2$ and $\dot{W}_{ext}$ relationship.

Parameter	Estimate	Est.Error	CI_{lower}	CI_{upper}	Effects
Intercept	−8.530	8.842	−27.898	7.197	Population-level effects
Mechanical power	0.081	0.018	0.035	0.111	Population-level effects
sd(Intercept)	15.234	7.535	3.123	32.769	Group-level effects
sd(mechanical power)	0.029	0.019	0.004	0.078	Group-level effects
cor(Intercept,mechanical power)	−0.492	0.428	−0.949	0.658	Group-level effects
sigma	2.410	0.346	1.853	3.204	Family specific parameters

Oxygen uptake and work rate relationship were largely studied and well described by Poole, Barstow, Gaesser, et al. 1994 and Gaesser and Poole 1996. Our results agreed with the literature with a strong and positive relationship between consumed metabolic energy and $\dot{W}_{ext}$ across the MAS test (Table 4.1). To end with metabolic and mechanical power measures, the net ME ($ME = 55 \pm 3\%$) also confirmed suitable values for a running task, as stated in the literature (Cavagna and Kaneko 1977; Kyröläinen, Komi, and Belli 1995; Willems, Cavagna, and Heglund 1995).

4.2.2.2. Ground Contact Time and Leg Stiffness

Force platform GCT and k_{leg} values according to the method of McMahon and Cheng 1990 (Equation (4.3)) are represented in Figure 4.3. GCT showed a strong and negative linear relationship with both speed ($R^2 = 0.96$, $[0.86, 0.98]$ 95% CI) and frequency ($R^2 = 0.93$, $[0.79, 0.96]$ 95% CI), recorded by the force platform and the motion capture system, respectively. Bayes factors reported a moderate evidence for the alternative hypothesis ($BF_{10} = 3.07$ and $BF_{10} = 4.02$). k_{leg} showed a strong and positive linear relationship with the stride frequency ($R^2 = 0.82$, $[0.42, 0.89]$ 95% CI, $BF_{10} = 16.12$). In contrast, k_{leg} did not significantly increase with speed ($R^2 = 0.40$, $[0, 0.64]$ 95% CI). The resulting Bayes factor ($BF_{10} = 0.22$) indicated an anecdotal evidence in favour of the null hypothesis.

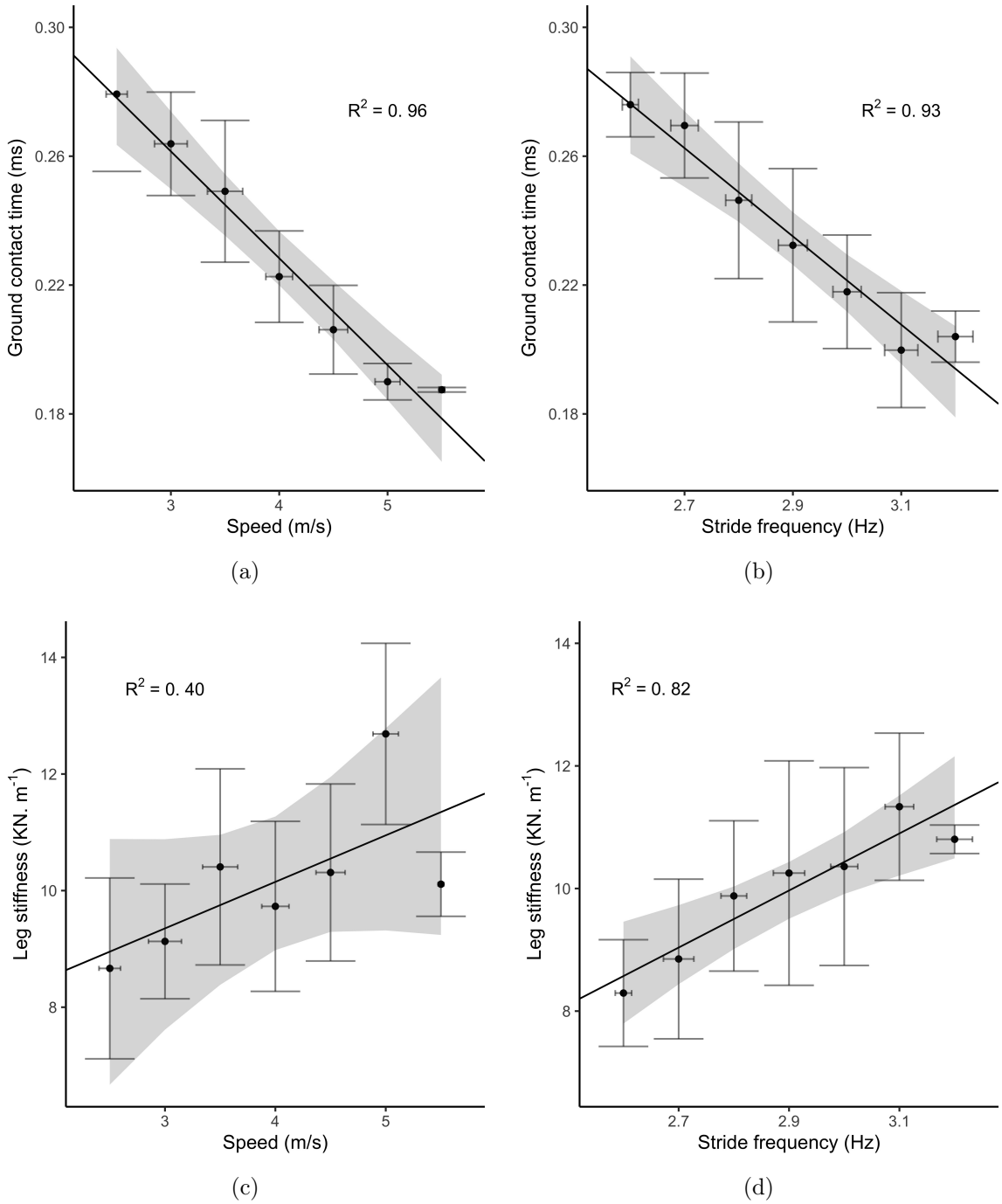


Figure 4.3. – Mechanical stride changes during the MAS test. The top plots (a, b) represent changes in GCT over speed and stride frequency respectively. The bottom plots (c, d) represent changes in k_{leg} over speed and stride frequency according to McMahon and Cheng 1990. In each figures, dots represent the group mean values, error bars the standard deviation in both x, y axes. The solid line is the regression line from the Bayesian linear model, surrounded by the 95% credible intervals.

Linear relationship (Figure 4.3) between both GCT, speed and stride frequency were in agreement with the literature (Carrard, Fontana, and Malatesta 2018; Farley and Gonzalez 1996). The leg stiffness calculated following McMahon's method (McMahon and Cheng 1990) also reported consistent results with previous author findings (Arampatzis, Bruggemann, and Metzler 1999; Farley and Gonzalez 1996; Morin, Dalleau, Kyröläinen, et al. 2005).

We considered that our reference measures were suitable for comparisons with the Stryd power meter based on these results.

4.2.3. Stryd and reference measures comparisons

4.2.3.1. Consumed Metabolic Energy and Power Output

Individual consumed metabolic energy and PO relationships are represented in Figure 4.4. We truncated the first part of the signal in order to remove the habituation period, where PO increases instantly and $\dot{V}O_2$ is shortly delayed. A strong and positive linear relationship between $\dot{V}O_2$ and PO for all participants was observed ($R^2 = 0.82$, $[0.81, 0.83]$ 95% CI, $BF_{10} > 100$). This results supported a valid relative measure of the Stryd PO from low speeds until MAS. Bayesian pair-wise correlation coefficient indicated a strong and positive correlation between both parameters for all participants ($r = 0.90$, $[0.89, 0.92]$ 95% CI, $BF_{10} > 100$).

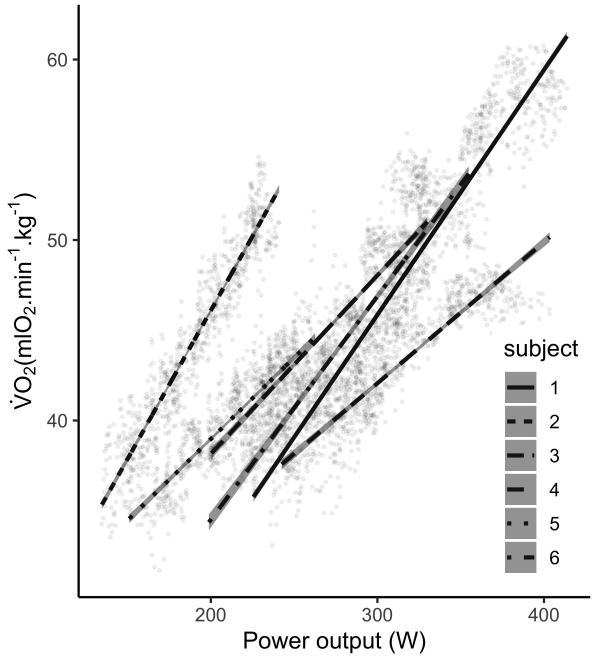


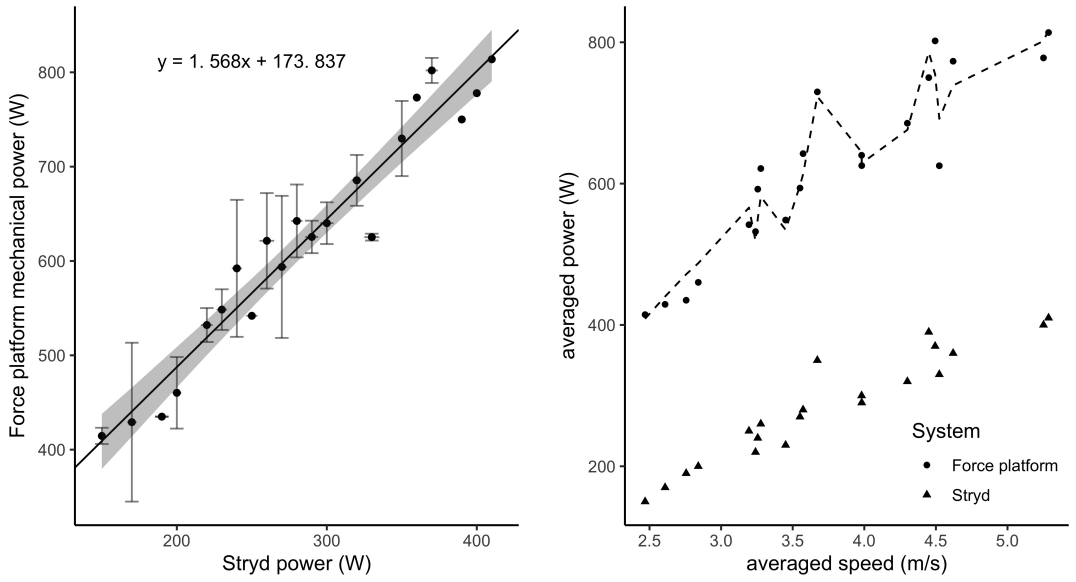
Figure 4.4. – $\dot{V}O_2$ – Stryd PO relationship during the incremental test. A strong and positive linear relationship was observed across participants. Lines represent each individual linear regression between $\dot{V}O_2$ and PO.

Individual regression slope and intercept differed between participants, according to heterogeneous body characteristics included in PO calculation (mainly body mass) and running performance level. Among the authors that assessed the metabolic demand and Stryd PO relationship, Aubry, Power, and Burr 2018 found a weak linear relationship between PO and $\dot{V}O_2$ but suffered from methodological flaws highlighted by Snyder, Mohrman, Williamson, et al. 2018.

Our results appear to be more consistent than Austin, Hokanson, McGinnis, et al. 2018 but we found lower correlations than Lara, Shearer, Coppi, et al. 2018 or Stryd own researches (Stryd 2017). It is important to point out the methodology employed when analysing $\dot{V}O_2$ from Cosmed k4b2 and Stryd power data. In the present study, the first part of the running exercise (less than one minute of exercise) was omitted due to a normal time delay for $\dot{V}O_2$ to increase. In contrast, PO increases instantaneously (Perrey 2010). Moreover, as the portable metabolic analyser provides a noisy signal (Duffield, Dawson, Pinnington, et al. 2004), filtering the data appears to be essential prior to any analysis. The fifth order moving averaged filter allowed analysis on smoothed data but still sensitive to the exercise conditions. We encourage such a process for $\dot{V}O_2$ analysis during an incremental running test or at least to report details about data processing.

4.2.3.2. Mechanical Power

A descriptive analysis of PO differences between both measurement systems indicated the greatest differences at highest speeds, suggesting a proportional error (normalized PO differences per participants varied from 38% to 60% between the two systems). By modelling the averaged PO across participants, a strong and linear relationship was observed over the MAS test (Figure 4.5a, $R^2 = 0.94$, $[0.91, 0.95]\,95\%\,\text{CI}$, $BF_{10} > 100$). However, the two systems were different regarding their absolute measures. To fix this issue, the linear model fitted on the averaged values of each participant measures could provide a correction function. In this study the estimated function was of the form $f(x) = ax + b$, with $a = 173.837$ and $b = 1.569$ (Figure 4.5b). In addition, the Bland-Altman analysis provided a representation of such differences in absolute values and a proportional error which raised with mechanical power increase, and so speed (see Figure 4.6a).



(a) (b)

Figure 4.5. – Comparison of PO estimated by the Stryd power meter and the force platform. The left plot (**a**) represents the strong positive relationship between the Stryd and the reference measures. The right plot (**b**) represents the averaged PO in response to speed, where the dotted line is the corrected Stryd PO (see text for details).

In addition to averaged power comparisons, LMM mean posterior distributions and credible intervals confirmed the underestimate of the Stryd PO ($\beta_1 = -305$, $[-324, -286]\,95\%\,\text{CI}$, $BF_{10} > 100$). Results are reported in Table 4.2. In agreement with the relationship of $\dot{W}_{ext}$ and speed, the posterior distribution of the speed

parameter also reported a positive effect on PO with a strong evidence ($BF_{10} = 15.38$). Stryd measure and speed interaction reported a negative effect. It suggests that the power meter measure is not homogeneous while speed raises (from 8 km/h to approximately 19 km/h). By comparing models with and without the interaction term, Bayes factor confirmed this effect with an extreme evidence ($BF_{10} > 100$). The ICC revealed a high correlation ($ICC = 0.79$, $[0.55, 0.96]$ 95% CI, estimated error = 0.11). Such a correlation indicated an important variance explained by the random intercept and slopes per subject, justifying the relevance of subject random terms in the present modelling.

Results found from the LMM (Table 4.2) and the descriptive analysis of differences between the Stryd power meter and the force platforms converged, suggesting a proportional error of the Stryd PO estimate with the increase of speed (Figure 4.6a). The underestimate of the power meter measure could be accounted for by the apparent ME used in the power calculation. Apparent efficiency represents the ratio between mechanical energy and metabolic energy and is modulated by elastic storage and recoil from the eccentric to the concentric phase. It has been measured during running, walking and cycling (Asmussen and Bonde-Petersen 1974; Cavagna and Kaneko 1977; Kyröläinen, Komi, and Belli 1995; Willems, Cavagna, and Heglund 1995; Zacks 1973). Researchers found values up to 53% in running versus approximately 30% and 25% in walking and cycling respectively (we found similar values of 55%, as reported above). Running power meters are concerned with ME assessment, but the way it has been integrated into the PO estimation could explain large differences in absolute PO values. Stryd team mentioned in their white papers a gross ME equal to 25% that can be approached by elite runners (Stryd 2017). This ME value is about half of the apparent ME values reported by the aforementioned authors and the one we found. Such a difference could explain the underestimate of PO from the Stryd power meter when compared to the systems of reference. To tackle this absolute error, we proposed a linear correction function adjusting the power value to the appropriate scale (Figure 4.5b). The provided function is estimated from data of only six runners. Even though their heterogeneous aptitude and body mass varied, further studies including more participants could provide a more accurate PO correction for a wide range of runners.

4.2.3.3. Ground Contact Time

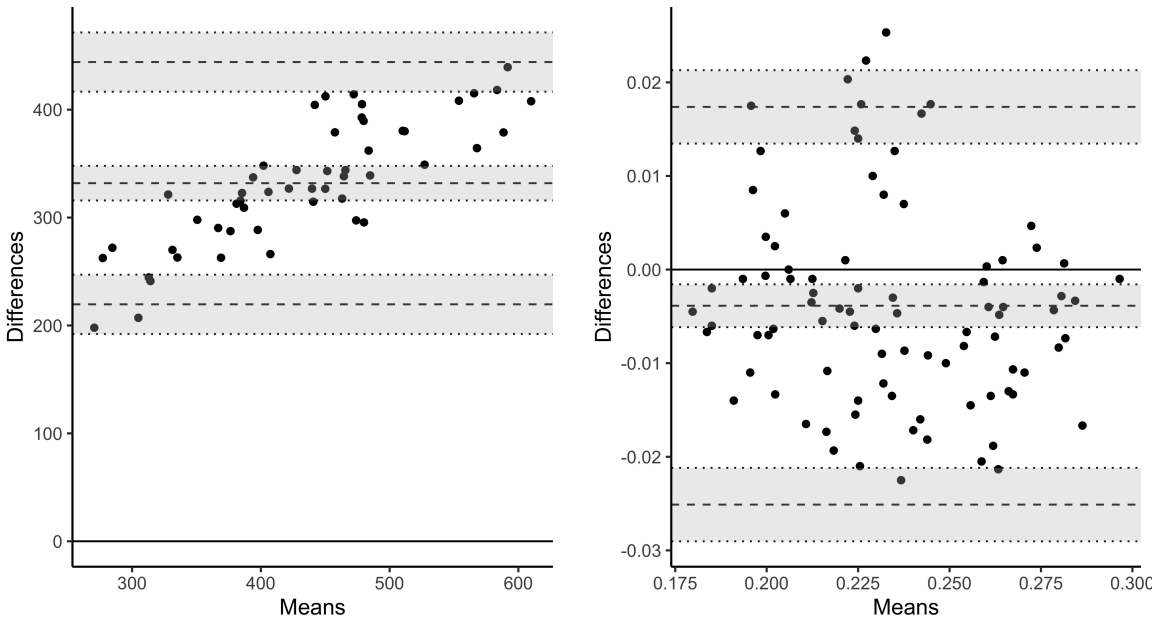
Results of the Bayesian LMM (Table 4.2) reported a negative although quasi-null posterior estimate of the Stryd parameter ($\beta_1 = -0.005$, $[-0.009, -0.002]$95%CI). It suggested that a small but negative effect of the Stryd device exists. However, Bayes factor reported an extreme evidence in favour of the null hypothesis ($BF_{10} = 0.008$). Hence, no significant differences were found between the systems of measurement. Posterior distribution of the speed parameter also reported a negative effect on GCT in agreement with our precedent results (see Figure 4.3). Bayes factor supported

the importance of this effect with an extreme evidence ($BF_{10} > 100$). However, not any effect were observed for the Stryd power meter and speed interaction. Bayes factor confirmed this result with an anecdotal evidence against the alternative hypothesis ($BF_{10} = 0.72$). Thus, GCT seemed to be measured with consistency across the whole range of speeds. In terms of by subject explained variance, the ICC reported a moderate correlation ($ICC = 0.55$, $[0.24, 0.89]$ 95% CI, estimated error = 0.18). Therefore, individual differences in GCT were lower than PO differences.

According to Garcı+a-Pinillos, Roche-Seruendo, Marcén-Cinca, et al. 2018, the Stryd system provides an underestimate and poor reliability for the GCT measure by comparing it to an OptoGait system. No such results were found in the present study, although an underestimate, negligible GCT measure was observed when compared to force platforms (Table 4.2, Figure 4.6b). A potential difference could be explained in the Stryd vertical GRFs modelling itself where the passive peak is missed (Stryd 2017). Moreover, as discussed by the authors, the relevance of the OptoGait system would not be as accurate as our system of reference (force platform), which could explain divergent findings.

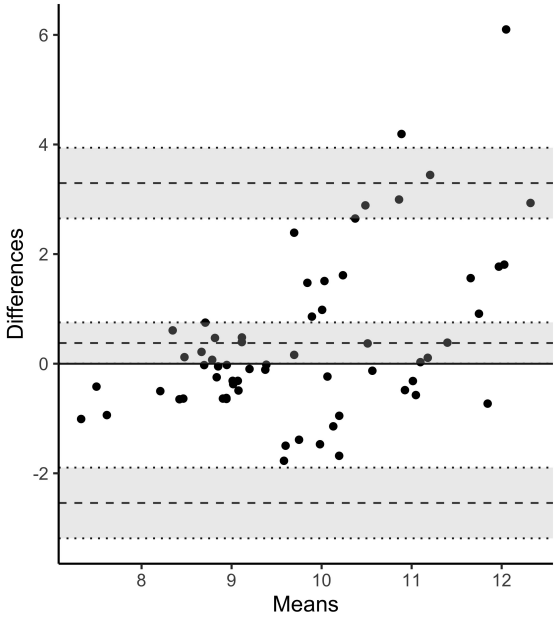
Table 4.2. – Bayesian linear mixed models parameters estimates (see Appendix H.2 for details about models.

Parameter	Estimate	Est.Error	CI_{lower}	CI_{upper}	BF_{10}	Effects	Measure
Intercept	568.401	38.163	490.765	644.830		Population-level effects	Mechanical power
Stryd	−304.952	9.817	−324.304	−285.993	>100	Population-level effects	Mechanical power
Speed	65.486	12.914	42.022	93.125	15.38	Population-level effects	Mechanical power
Stryd:speed interaction	−23.782	9.884	−43.192	−4.554	>100	Population-level effects	Mechanical power
sd(Intercept)	86.027	36.563	41.970	180.565		Group-level effects	Mechanical power
sd(speed)	17.260	18.408	0.438	66.867		Group-level effects	Mechanical power
cor(Intercept,speed)	0.349	0.516	−0.819	0.978		Group-level effects	Mechanical power
sigma	52.325	3.651	45.751	59.970		Family specific parameters	Mechanical power
Intercept	0.241	0.008	0.226	0.255		Population-level effects	Contact time
Stryd	−0.005	0.002	−0.009	−0.002	0.008	Population-level effects	Contact time
Speed	−0.034	0.012	−0.058	−0.011	>100	Population-level effects	Contact time
Stryd:speed interaction	0.000	0.002	−0.003	0.004	0.72	Population-level effects	Contact time
sd(Intercept)	0.016	0.009	0.007	0.039		Group-level effects	Contact time
sd(speed)	0.025	0.015	0.009	0.063		Group-level effects	Contact time
cor(Intercept,speed)	−0.216	0.410	−0.862	0.647		Group-level effects	Contact time
sigma	0.010	0.001	0.009	0.012		Family specific parameters	Contact time
Intercept	8.574	0.980	6.680	10.571		Population-level effects	Leg stiffness
Stryd	−0.602	0.893	−2.334	1.154	0.007	Population-level effects	Leg stiffness
Speed	0.394	0.240	−0.099	0.865	0.012	Population-level effects	Leg stiffness
Stryd:lap interaction	0.063	0.244	−0.418	0.534	0.020	Population-level effects	Leg stiffness
sd(Intercept)	1.427	0.959	0.119	3.830		Group-level effects	Leg stiffness
sd(speed)	0.284	0.255	0.010	0.940		Group-level effects	Leg stiffness
cor(Intercept,speed)	−0.105	0.572	−0.953	0.927		Group-level effects	Leg stiffness
sigma	0.972	0.063	0.857	1.106		Family specific parameters	Leg stiffness



(a) mechanical power

(b) Ground contact time



(c) Leg stiffness

Figure 4.6. – Bland–Altman plots for comparison of measurements between the force platforms (reference) and the power meter. Mean bias (middle dashed line), lower and upper limits of agreement (dashed lines) and their 95% confidence interval areas are represented.

4.2.3.4. Leg Spring Stiffness

For the last measure of interest, any difference were found between devices over the LSS measurement ($\beta_1 = -0.602$, $[-2.334, 1.154]95\%$CI) as reported in Table 4.2 and presented in Figure 4.6c. An extreme evidence against the alternative hypothesis supported this result ($BF_{10} = 0.007$). The speed did not appear to impact the LSS measure which remained quite constant. Bayes factor supported this result with a very strong evidence against the alternative hypothesis ($BF_{10} => 0.012$). Finally, the posterior distribution of the LSS parameter indicated that LSS was measured with consistency across the range of speed. A very strong evidence against the alternative hypothesis supported the negligible effect in modelling ($BF_{10} => 0.020$). Nonetheless, by between subject variability showed a moderate correlation ($ICC = 0.65$, $[0.34, 0.93]$ 95% CI, estimated error = 0.16) revealing differences in individual baseline levels as well as responses to speed increment.

The relationship found in the present study between leg stiffness and stride frequency was in agreement with the literature (Farley and Gonzalez 1996). In addition, Morin, Samozino, Zameziati, et al. 2007 found that 90% to 96% of changes in leg stiffness was accounted for by changes in GCT. In contrast, step frequency indirectly influenced leg stiffness through its relationship with GCT. On the other hand, the leg stiffness represents the lower-limb resistance to deformation and reflects in some way the elastic energy storage and recoil. Thus, leg stiffness has been considered as a kinetic factor related to running economy (REc) by authors (Moore 2016; Slawinski, Heubert, Quievre, et al. 2008). REc is recognised to be one of the main determinants of the endurance running performance (Foster and Lucia 2007). Thereby, LSS could be a relevant training parameter in which the runner should major it by mainly shortening the GCT (in addition to resistance and specific training). However, further studies involving changes in GCT, stride frequency and kinematic factors (e.g., angle of attack) for a given speed would assess whether the Stryd power meter is sensitive enough to correctly estimate REc through the LSS.

4.2.4. Conclusion

The power meter provided acceptable measures of GCT and LSS over the test, but interrogations persist about absolute values of PO. Nonetheless, the Stryd power meter can be a valuable tool to quantify the intensity during sub-maximal running. By correcting absolute PO values of the Stryd power meter to get true $\dot{W}_{ext}$ it allows the runner to monitor training loads (*e.g.* through the external work) and performances across sessions. Recreational runners interested in health rather than performance can also profit of these measures by practising safely. Further analyses remain necessary to assess the power meter's validity at higher speeds (maximal and supra-maximal), with direction changes, non-linear accelerations and slope, and clearly define which power is measured.

4.2.5. Take-home message

— The Stryd power meter provided valid measures of GCT and LSS but underestimated the true values of $\dot{W}_{ext}$.
— Interrogations persist since PO reported by the system differs from conventional $\dot{W}_{met}$, W and $\dot{W}_{ext}$. However, the system seems reliable and could therefore be used for training programming and monitoring purposes.

The following sections present some tools developed alongside collaborators.

4.3. *RunningPerfApp*: A *shiny* web-application for long distance runners

Shiny web-applications are developed from the *Shiny* open-source *R* package, which provides a powerful web framework or building interactive applications. *RunningPerfApp* (Imbach 2019) has been developed for long-distance runners who were preparing for a marathon. Over a three-month study period, we followed six runners equipped with the *Stryd* running power meter (Boulder, Colorado, United States), smartwatches (Garmin Fenix 5, Suunto Spartan) and HR monitors.

4.3.1. Monitoring athlete status

After having assessed the validity of the power meter (Imbach, Candau, Chailan, et al. 2020), we provided to runners key insights about their athletic status (see Figure 4.7).

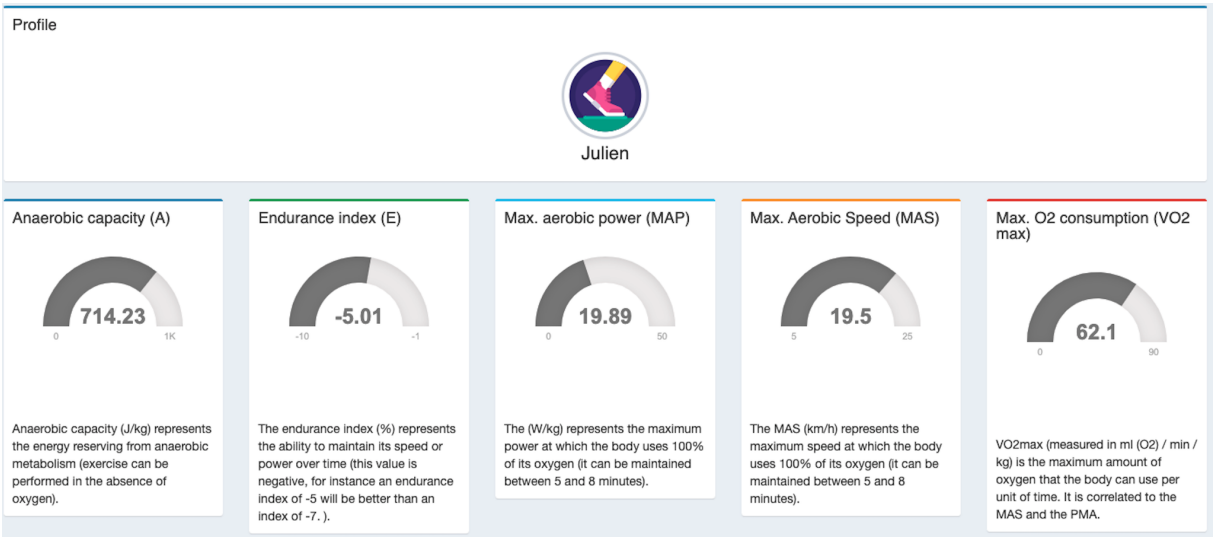
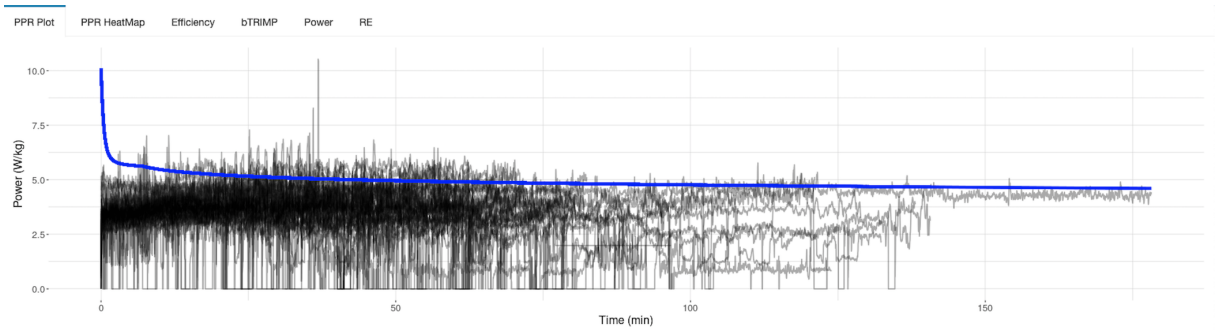


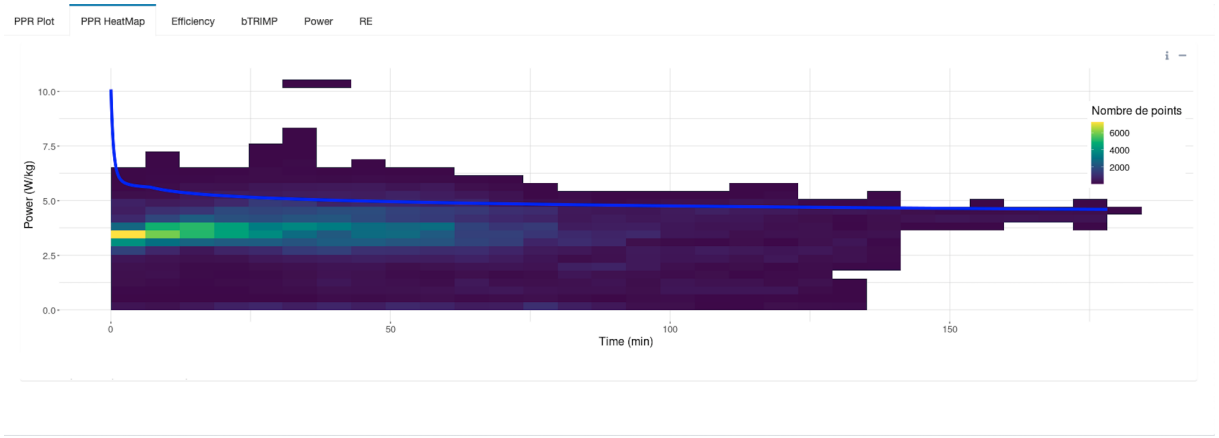
Figure 4.7. – Profile for a given runner.

In the first instance, runners performed time trials on various distances (800m, 2400m, 3000m, 5000m, 10000m) in order to model their physiological capacities, according to the mathematical model of Péronnet and Thibault 1989. Since they were preparing for long-distance runs in which the contribution of the anaerobic metabolism remain small, they did not perform shorter time trials (*e.g.* 100m, 200m, 400m). Also, runners were asked to perform alternately 3000m and 5000m every two weeks in order to evaluate the changes in time trial performances as well as modelled physiological parameters highlighted in Figure 4.7.

Each training session was graphically represented along with the record power profile (RPP) which should theoretically represent their performance potential, based on the modelled physiological parameters (see Figures 4.8a and 4.8b).



(a)



(b)

Figure 4.8. – Record power profile modelled according to Péronnet and Thibault 1989 for a given runner. In (a), black lines represent training sessions the power measured at each training session and the blue line is the record power profile estimated from time-trials. In (b), a heat map is displayed in order to represent distributions of the running power according to the time.

In addition, we provided a daily performance index, namely an efficiency described as by the ratio between relative mechanical and metabolic indicators of intensity. Hence, we define efficiency at day i such as

$$Efficiency_i = \frac{1}{N}\sum_{n=1}^{N}\frac{\dot{W}_{i,t}^{\text{ext}^*}}{\dot{W}_{i,t}^{\text{int}^*}}\,, \quad \text{with}$$
$$\dot{W}_{i,t}^{\text{ext}^*} = \frac{\dot{W}_{i,t}^{\text{ext}}}{\text{RPP}_t^{\text{ext}}} \quad \text{and}$$
$$\dot{W}_{i,t}^{\text{int}^*} = \frac{\text{bTRIMP}_{i,t}}{max(\text{bTRIMP})_t}\,.$$

with $\dot{W}_{i,t}^{\text{ext}}$ being the $\dot{W}_{ext}$ at day i and time t, $\text{RPP}_t^{\text{ext}}$ the value of the RPP at time t, $\text{bTRIMP}_{i,t}$ is the bTRIMP value at day i and time t. Accordingly, a high efficiency describes a low contribution of the internal mechanisms (*i.e.* calculated from HR) for running at a given power output, and inversely. An example of such performance criterion is provided in Figure 4.9.

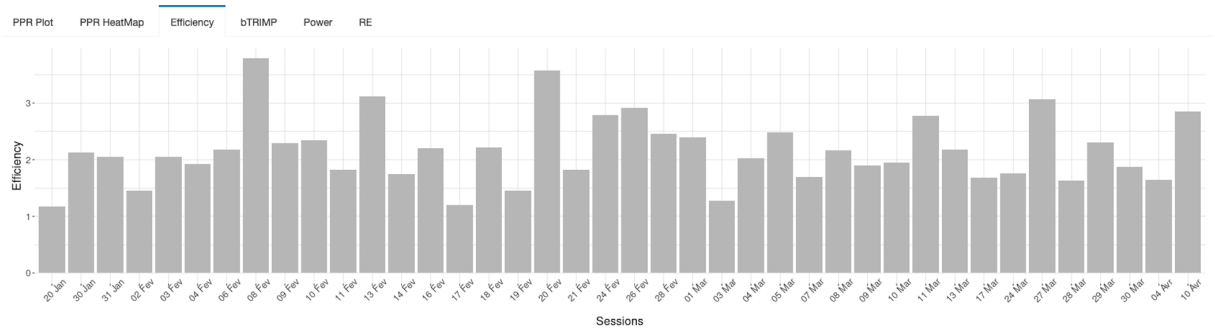


Figure 4.9. – Daily performances of a given runner over the study period.

Finally, we displayed a predictive module based on a linear extrapolation of the physiological model (Péronnet and Thibault 1989). A given target running time returns the optimal running power to be held all along the race.

4.4. Team management and data visualisation with SAP analytics cloud

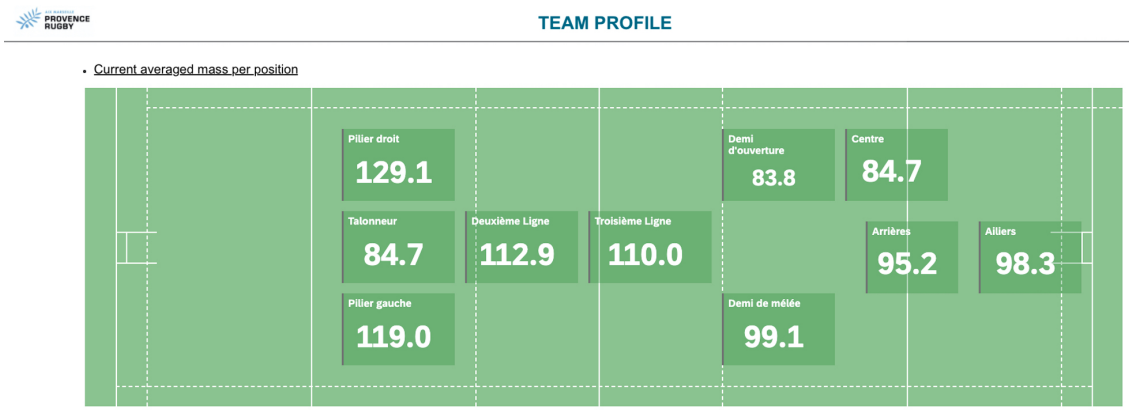
SAP analytics cloud (a trademark SAP, Germany) is a cloud solution for business intelligence, data visualisation and planning. It provides great compliance, security, and data protection practices and allows for easy implementation of predictive analysis modules.

In parallel to the academic work presented in the previous chapters, we worked alongside a professional Rugby team, *Provence Rugby* (second professional division,

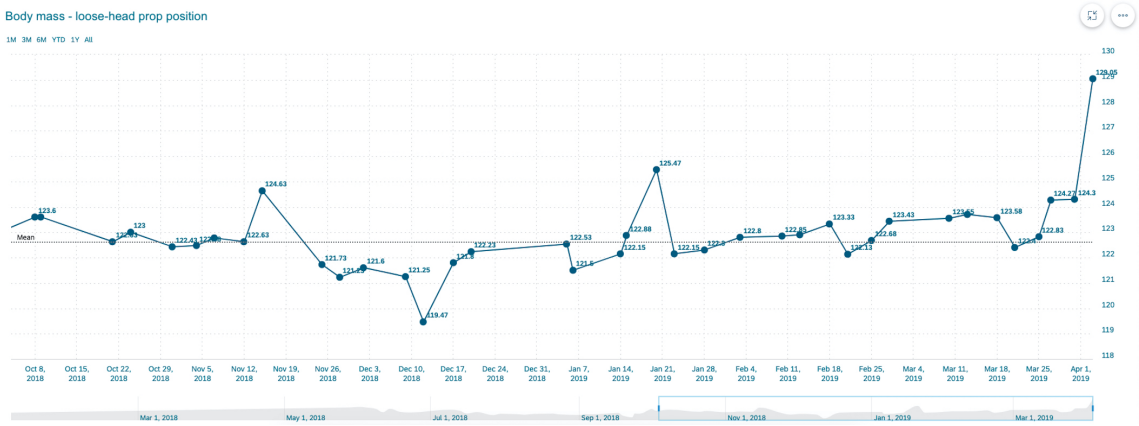
Pro D2) for the 2018-2019 season. The collaboration allowed us to gather GPS statistics (McLloyd, TIPA, France), well-being indicators (Jeffries, Wallace, Coutts, et al. 2020), anthropometric measures and physical performances. Some examples of the developed dashboard are presented hereafter.

4.4.1. Monitoring team and players body mass

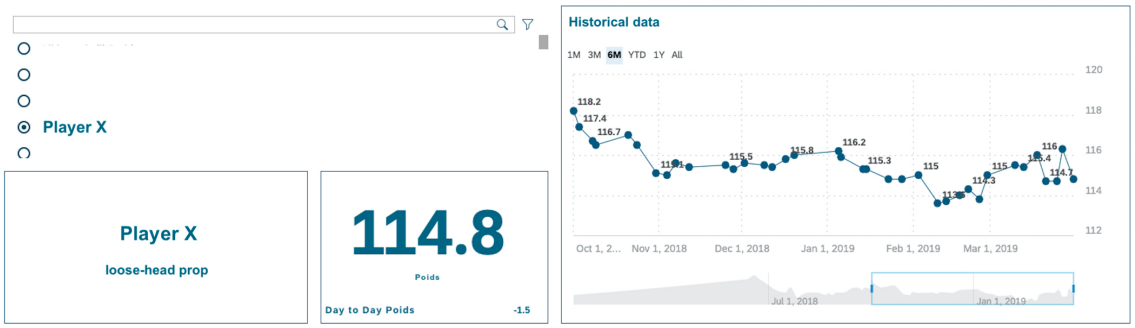
Body mass is a basic parameter to monitor in Rugby union. It informs on player changes over the season and is rendered more or less significant depending on the position played. A team-report and player-report is provided through Figures 4.10a, 4.10b and 4.10c. Note that for confidentiality reasons, body mass data were simulated.



(a)



(b)

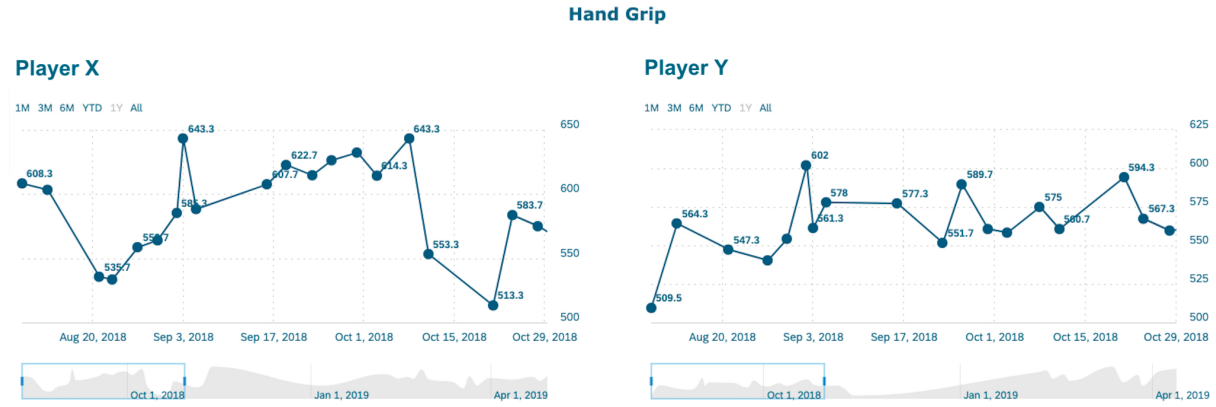


(c)

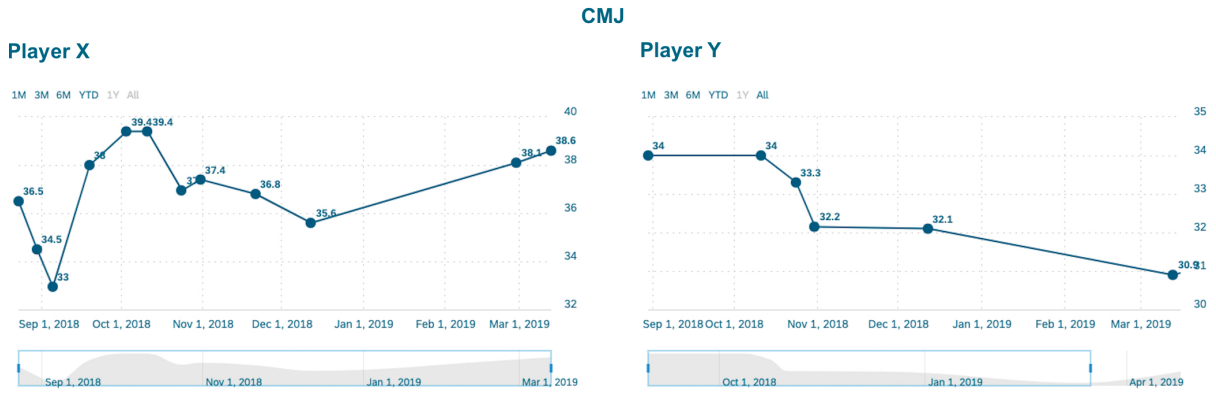
Figure 4.10. – Summary of body mass measures with (a) average per position, (b) the progression of averaged mass of the loose-head prop position (c) a body mass monitoring of a given player playing at loose-head prop position.

4.4.2. Monitoring player performances

In Figures 4.11a and 4.11b, we presented two performances indicators that are regularly performed for readiness and muscular performances evaluation. By way of illustration, one may find a significant decay in hand grip and CMJ performance for *player X* during October. Considering the player not being injured, it might represent an alarm signal for coaches. In addition, two opposite trends of CMJ performances are observed among both represented players from January 2019. Such information should inform and guide technical and athletic coaches in their training prescription and team management.



(a)



(b)

Figure 4.11. – Comparison of (a) hand grip and (b) CMJ performances between players.

4.4.3. Monitoring training from global positioning system data

Wearing GPS during each training and game sessions allow for monitoring exercise through a quantitative representation of volume and intensity parameters. Figure 4.12 represents the number of accelerations performed at different levels as well as distance covered at various velocities.

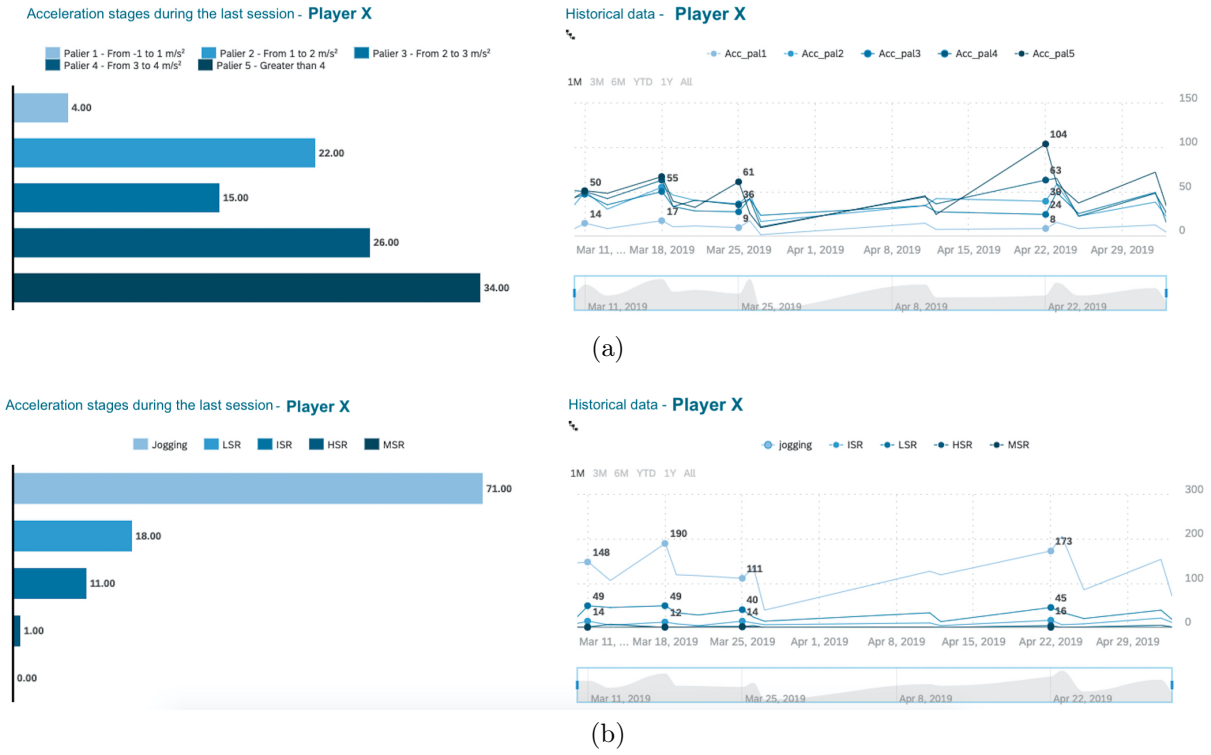


Figure 4.12. – Example of training monitoring through some commercial GPS features with (a) the evolution of acceleration-related features and (b) the evolution of velocity-related features over the last month.

4.4.4. Monitoring Well-being

A final example shows the basic usage of well-being monitoring. Since well-being estimation methods are numerous (Jeffries, Wallace, Coutts, et al. 2020), we provided a few items of interest in Figure 4.13b. A summarises of day-to-day changes is generally useful for quickly adjusting training sessions, for instance, in case of severe muscular soreness or fatigue.

Figure 4.13. – Well-being dashboard with (a) a summary of well-being status for a given player, and (b) the evolution of each well-being item.

Conclusion

In this thesis, we provided comprehensive modelling of athletic performance in athletes. First, we exposed a state of the art about the usual methods of training loads quantification and the models of training effects so far known.

Then, we concentrated on the quantification of TL in resistance training and proposed a new method based on physiological responses to exercise. This method is intended to be applicable in any context of RT for athlete monitoring purposes.

After that, we proceeded to two predictive applications in individual and collective sports, respectively. In the first, we presented a robust methodology relying on model generalisation and compared a dose-response model framework to statistics and machine-learning models. It was found that multivariate approaches with explanatory variables from various sources were better suitable for predicting athletic performance outcomes. In addition, a comparison between individually computed models and models that account for all data indicated that there were no significant advantages of fitting models per athlete in a small sample size context.

A second application was carried out in Football. We focused on modelling athletic performance directly from global positioning system measurements without requiring any specific evaluations. Hence, we attempted to make the most of GPS systems for modelling such performance with GPS data only. In line with the previous work and, more generally, the literature, we suggested that athletic performance is multidimensional and governed by several determinants from technical, tactical, psychological and other performance-related factors. Also, we highlighted the potential richness of raw data from GPS wearable sensors for athletic performance and injury modelling purposes.

Finally, we developed athlete management and decision-support systems for amateurs and professional athletes. Concurrently and still regarding athlete monitoring problems, we validated a promising measurement system for long-distance runners.

The work carried out during the thesis lets us state that athletic performance modelling is a complex problem that requires a systemic approach. Consequently, system models from control theory built on strong expert hypotheses would deserve further extensions and integration within machine-learning approaches for optimising training programs for performance.

Bibliography

Abdessemed, D, P Duche, C Hautier, et al. (1999). "Effect of recovery duration on muscular power and blood lactate during the bench press exercise". In: *Int. J Sports Med.* 20.06, pp. 368–373 (cit. on p. 50).

Achten, Juul and Asker E Jeukendrup (2003). "Heart rate monitoring". In: *Sports Med.* 33.7, pp. 517–538 (cit. on pp. 37, 38).

Agostinho, Marcus F, Antony G Philippe, Gilvan S Marcolino, et al. (2015). "Perceived training intensity and performance changes quantification in judo". In: *J. Strength Cond. Res.* 29.6, pp. 1570–1577 (cit. on p. 57).

Ahtiainen, Juha P, Arto Pakarinen, William J Kraemer, et al. (2003). "Acute hormonal and neuromuscular responses and recovery to forced vs. maximum repetitions multiple resistance exercises". In: *Int. J. Sports Med.* 24.06, pp. 410–418 (cit. on p. 49).

Akenhead, Richard and George P Nassis (2016). "Training load and player monitoring in high-level football: current practice and perceptions". In: *Int. J. Sports Physiol. Perform.* 11.5, pp. 587–593 (cit. on p. 166).

Allen, DG and H Westerblad (2001). "Role of phosphate and calcium stores in muscle fatigue". In: *J. Physiol.* 536.3, pp. 657–665 (cit. on p. 115).

Almuzaini, Khalid S, Jeffrey A Potteiger, and Samuel B Green (1998). "Effects of split exercise sessions on excess postexercise oxygen consumption and resting metabolic rate". In: *Can. J. Appl. Physiol.* 23.5, pp. 433–443 (cit. on p. 42).

Andersen, Lars L, Jesper L Andersen, Mette K Zebis, et al. (2010). "Early and late rate of force development: differential adaptive responses to resistance training?" In: *Scand. J. Med. Sci. Sports* 20.1, e162–e169 (cit. on p. 96).

Arampatzis, Adamantios, Gert-Peter Bruggemann, and Verena Metzler (1999). "The effect of speed on leg stiffness and joint kinetics in human running". In: *J. Biomech.* 32.12, pp. 1349–1353 (cit. on p. 200).

Arney, Blaine E, Reese Glover, Andrea Fusco, et al. (2019). "Comparison of RPE (rating of perceived exertion) scales for session RPE". In: *Int. J. Sports Physiol. Perform.* 14.7, pp. 994–996 (cit. on p. 57).

Asmussen, Erling and Flemming Bonde-Petersen (1974). "Storage of elastic energy in skeletal muscles in man". In: *Acta Physiol. Scand.* 91.3, pp. 385–392 (cit. on p. 203).

Assimakopoulos, Vassilis and Konstantinos Nikolopoulos (2000). "The theta model: a decomposition approach to forecasting". In: *Int. J. Forecast.* 16.4, pp. 521–530 (cit. on p. 176).

Atzmueller, Martin (2015). "Subgroup discovery". In: *Data. Min. Knowl. Disc.* 5.1, pp. 35–49 (cit. on p. 84).

Aubry, Rachel L, Geoff A Power, and Jamie F Burr (2018). "An assessment of running power as a training metric for elite and recreational runners". In: *J. Strength Cond. Res.* 32.8, pp. 2258–2264 (cit. on p. 201).

Austin, Casey L, James F Hokanson, Peter M McGinnis, et al. (2018). "The relationship between running power and running economy in well-trained distance runners". In: *Sports* 6.4, p. 142 (cit. on p. 201).

Avalos, Marta, Philippe Hellard, and Jean-Claude Chatard (2003). "Modeling the training-performance relationship using a mixed model in elite swimmers". In: *Med. Sci. Sports Exerc.* 35.5, p. 838 (cit. on pp. 64, 82, 83, 85, 154, 155).

Baird, Marianne F, Scott M Graham, Julien S Baker, et al. (2012). "Creatine-kinase-and exercise-related muscle damage implications for muscle performance and recovery". In: *J Nutr Metab.* 2012 (cit. on p. 48).

Baldari, Carlo, Valerio Bonavolontà, Gian Pietro Emerenziani, et al. (2009). "Accuracy, reliability, linearity of Accutrend and Lactate Pro versus EBIO plus analyzer". In: *Eur. J. Appl. Physiol.* 107.1, pp. 105–111 (cit. on p. 95).

Balsalobre-Fernández, Carlos and Lorena Torres-Ronda (2021). "The Implementation of Velocity-Based Training Paradigm for Team Sports: Framework, Technologies, Practical Recommendations and Challenges". In: *Sports* 9.4, p. 47 (cit. on p. 51).

Banister, Eric W, Thomas W Calvert, Margaret V Savage, et al. (1975). "A systems model of training for athletic performance". In: *Aust. J. Sports Med.* 7.3, pp. 57–61 (cit. on pp. 32, 59, 64, 67, 73, 154, 161, 163).

Banister, Eric W, James B Carter, and Peter C Zarkadas (1999). "Training theory and taper: validation in triathlon athletes". In: *Eur. J. Appl. Physiol. Occup. Physiol.* 79.2, pp. 182–191 (cit. on p. 83).

Banister, Eric W, Pat Good, Geoffrey Holman, et al. (1986). "Modeling the training response in athletes". In: *Sport and elite performers*. Vol. 3. Human Kinetics, pp. 7–23 (cit. on p. 260).

Banister, Eric W and Cindy L Hamilton (1985). "Variations in iron status with fatigue modelled from training in female distance runners". In: *Eur. J. Appl. Physiol. Occup. Physiol.* 54.1, pp. 16–23 (cit. on pp. 38, 39, 64, 85, 126).

Banister, EW and JR Fitz-Clarke (1993). "Plasticity of response to equal quantities of endurance training separated by non-training in humans". In: *J. Therm. Biol.* 18.5-6, pp. 587–597 (cit. on p. 64).

Banister, EW, H Green, J McDougall, et al. (1991). "Physiological testing of elite athletes". In: Champaign: IL: Human Kinetics, pp. 403–424 (cit. on p. 64).

Banyard, Harry G, Kazunori Nosaka, Alex D Vernon, et al. (2018). "The reliability of individualized load–velocity profiles". In: *Int. J. Sports Physiol. Perform.* 13.6, pp. 763–769 (cit. on p. 51).

Banyard, Harry G, Ken Nosaka, Kimitake Sato, et al. (2017). "Validity of various methods for determining velocity, force, and power in the back squat". In: *Int. J. Sports Physiol. Perform.* 12.9, pp. 1170–1176 (cit. on p. 51).

Banyard, Harry G, James J Tufano, Jose Delgado, et al. (2018). "Comparison of velocity-based and traditional 1RM-percent-based prescription on acute kinetic and kinematic variables". In: *Int. J. Sports Physiol. Perform.* (cit. on p. 51).

Barr, Dale J, Roger Levy, Christoph Scheepers, et al. (2013). "Random effects structure for confirmatory hypothesis testing: Keep it maximal". In: *J. Mem. Lang.* 68.3, pp. 255–278 (cit. on p. 265).

Barrett, Steve, Adrian Midgley, and Ric Lovell (2014). "PlayerLoad™: reliability, convergent validity, and influence of unit position during treadmill running". In: *Int. J. Sports Physiol. Perform.* 9.6, pp. 945–952 (cit. on p. 53).

Barstow, Thomas J (2019). "Understanding near infrared spectroscopy and its application to skeletal muscle research". In: *J. Appl. Physiol.* 126.5, pp. 1360–1376 (cit. on p. 96).

Barstow, THOMAS J, RICHARD Casaburi, and KARLMAN Wasserman (1993). "O2 uptake kinetics and the O2 deficit as related to exercise intensity and blood lactate". In: *J. Appl. Physiol.* 75.2, pp. 755–762 (cit. on p. 37).

Barstow, Thomas J and Paul A Molé (1991). "Linear and nonlinear characteristics of oxygen uptake kinetics during heavy exercise". In: *J. Appl. Physiol.* 71.6, pp. 2099–2106 (cit. on p. 37).

Bartlett, R (1997). "The use and abuse of statistics in sports and exercise sciences." In: *J. Sports Sci.* 15.1, pp. 1–2 (cit. on p. 63).

Bates, Douglas M and Donald G Watts (1988). *Nonlinear regression analysis and its applications*. Vol. 2. Wiley New York (cit. on p. 64).

Bazyler, Caleb D, Heather A Abbott, Christopher R Bellon, et al. (2015). "Strength training for endurance athletes: theory to practice". In: *Strength Cond. J.* 37.2, pp. 1–12 (cit. on pp. 32, 59, 64).

Bergmeir, Christoph and José M Benı+tez (2012). "On the use of cross-validation for time series predictor evaluation". In: *Inf. Sci.* 191, pp. 192–213 (cit. on pp. 141, 159).

Bergmeir, Christoph, Mauro Costantini, and José M Benı+tez (2014). "On the usefulness of cross-validation for directional forecast evaluation". In: *Comput. Stat. Data Anal.* 76, pp. 132–143 (cit. on p. 141).

Bergmeir, Christoph, Rob J Hyndman, and Bonsoo Koo (2018). "A note on the validity of cross-validation for evaluating autoregressive time series prediction". In: *Comput. Stat. Data Anal.* 120, pp. 70–83 (cit. on p. 141).

Bickel, Peter J, Bo Li, Alexandre B Tsybakov, et al. (2006). "Regularization in statistics". In: *Test* 15.2, pp. 271–344 (cit. on p. 83).

Billat, VL, DW Hill, J Pinoteau, et al. (1996). "Effect of protocol on determination of velocity at VO2 max and on its time to exhaustion". In: *Arch. Physiol. Biochem.* 104.3, pp. 313–321 (cit. on p. 193).

Biolo, Gianni, Sergio P Maggi, Bradley D Williams, et al. (1995). "Increased rates of muscle protein turnover and amino acid transport after resistance exercise in humans". In: *Am. J. Physiol.* 268.3, E514–E520 (cit. on p. 47).

Bird, Stephen P, Kyle M Tarpenning, and Frank E Marino (2005). "Designing resistance training programmes to enhance muscular fitness". In: *Sports Med.* 35.10, pp. 841–851 (cit. on p. 50).

Bishop, Christopher M (2006). "Pattern recognition". In: *Mach. Learn.* 128.9 (cit. on p. 138).

Blickhan, Reinhard (1989). "The spring-mass model for running and hopping". In: *J. Biomech.* 22.11-12, pp. 1217–1227 (cit. on p. 265).

Bobbert, Maarten F (2012). "Why is the force-velocity relationship in leg press tasks quasi-linear rather than hyperbolic?" In: *J. Appl. Physiol.* 112.12, pp. 1975–1983 (cit. on p. 135).

Bodor, Marko (1999). "Mechanomyographic and electromyographic muscle responses are related to power". In: *Muscle Nerve.* 22.5, pp. 649–650 (cit. on p. 120).

Bompa, Tudor O and G Gregory Haff (2009). *Periodization: Theory and methodology of training.[5-th Edition].* Champaign, IL, USA: Human Kinetics (cit. on p. 53).

Bond, Colin W, Emily M Willaert, and Benjamin C Noonan (2017). "Comparison of three timing systems: reliability and best practice recommendations in timing short-duration sprints". In: *J. Strength Cond. Res.* 31.4, pp. 1062–1071 (cit. on p. 144).

Bond, Colin W, Emily M Willaert, Kyle E Rudningen, et al. (2017). "Reliability of three timing systems used to time short on ice-skating sprints in ice hockey players". In: *J. Strength Cond. Res.* 31.12, pp. 3279–3286 (cit. on p. 144).

Borg, Gunnar (1970). "Perceived exertion as an indicator of somatic stress." In: *Scand. J. Rehabil. Med.* (cit. on p. 54).

– (1985). *An introduction to Borg's RPE-scale.* Mouvement Publications (cit. on p. 54).

– (1990). "Psychophysical scaling with applications in physical work and the perception of exertion". In: *Scand. J. Work Environ. Health.*, pp. 55–58 (cit. on pp. 55, 56).

– (1998). *Borg's perceived exertion and pain scales.* Human kinetics (cit. on pp. 54–56).

Borg, Gunnar and E Borg (2010). *The Bord CR Scales® folder.* Hasselby, Sweden, Borg Perception. (cit. on pp. 55, 56).

Boroujerdi, Saeed Sadeqi and Rahman Rahimi (2008). "Acute GH and IGF-I responses to short vs. long rest period between sets during forced repetitions resistance training system". In: *S. Afr. J. Res. Sport Ph.* 30.2, pp. 31–38 (cit. on p. 50).

Borresen, Jill and Michael I Lambert (2008). "Quantifying training load: a comparison of subjective and objective methods". In: *int. J. Sports Physiol. Perform.* 3.1, pp. 16–30 (cit. on p. 57).

Borresen, Jill and Michael Ian Lambert (2009). "The quantification of training load, the training response and the effect on performance". In: *Sports Med.* 39.9, pp. 779–795 (cit. on pp. 37–40, 56, 57).

Børsheim, Elisabet and Roald Bahr (2003). "Effect of exercise intensity, duration and mode on post-exercise oxygen consumption". In: *Sports Med.* 33.14, pp. 1037–1060 (cit. on pp. 42–44).

Bottaro, Martim, Breno Martins, Paulo Gentil, et al. (2009). "Effects of rest duration between sets of resistance training on acute hormonal responses in trained women". In: *J. Sci. Med. Sport.* 12.1, pp. 73–78 (cit. on pp. 47, 50).

Bouchard, Claude and Tuomo Rankinen (2001). "Individual differences in response to regular physical activity". In: *Med. Sci. Sports Exerc.* 33.6; SUPP, S446–S451 (cit. on p. 42).

Bourdon, Pitre C, Marco Cardinale, Andrew Murray, et al. (2017). "Monitoring athlete training loads: consensus statement". In: *Int. J. Sports Physiol. Perform.* 12.s2, S2–161 (cit. on pp. 57, 85, 166).

Boyd, Luke J, Kevin Ball, and Robert J Aughey (2013). "Quantifying external load in Australian football matches and training using accelerometers". In: *Int. J. Sports Physiol. Perform.* 8.1, pp. 44–51 (cit. on p. 53).

Brisswalter, Jeanick and Marcus P Tartaruga (2014). "Comparison of COSMED'S FitMate™ and K4b2 metabolic systems reliability during graded cycling exercise". In: *Sand. J. Clin. Lab. Invest.* 74.8, pp. 722–724 (cit. on pp. 94, 194).

Brzycki, Matt (1993). "Strength testing—predicting a one-rep max from reps-to-fatigue". In: *J. Phys. Educ. Recreat. Dance* 64.1, pp. 88–90 (cit. on p. 49).

Buchheit, Martin, Hani Al Haddad, Ben M Simpson, et al. (2014). "Monitoring accelerations with GPS in football: time to slow down?" In: *Int. J. Sports Physiol. Perform.* 9.3, pp. 442–445 (cit. on p. 166).

Bunc, Václav, Jan Heller, and Jirı+ Leso (1988). "Kinetics of heart rate responses to exercise". In: *J. Sports. Sci.* 6.1, pp. 39–48 (cit. on pp. 37, 38).

Burd, Nicholas A, Richard J Andrews, Daniel WD West, et al. (2012). "Muscle time under tension during resistance exercise stimulates differential muscle protein sub-fractional synthetic responses in men". In: *J. Physiol.* 590.2, pp. 351–362 (cit. on p. 51).

Burgess, Darren J (2017). "The research doesn't always apply: practical solutions to evidence-based training-load monitoring in elite team sports". In: *Int. J. Sports Physiol. Perform.* 12.s2, S2–136 (cit. on p. 59).

Burleson, Jr MA, HS O'Bryant, MH Stone, et al. (1998). "Effect of weight training exercise and treadmill exercise on post-exercise oxygen consumption". In: *Med. Sci. Sports Exerc.* 30.4, pp. 518–22 (cit. on p. 43).

Busso, Thierry (2003). "Variable dose-response relationship between exercise training and performance". In: *Med. Sci. Sports Exerc.* 35.7, pp. 1188–1195 (cit. on pp. 61, 66, 68, 69, 73, 80, 85, 144, 145, 149, 162).

Busso, Thierry (2017). "From an indirect response pharmacodynamic model towards a secondary signal model of dose-response relationship between exercise training and physical performance". In: *Sci. Rep.* 7.1, pp. 1–11 (cit. on pp. 78–80, 256).

Busso, Thierry, Henri Benoit, Regis Bonnefoy, et al. (2002). "Effects of training frequency on the dynamics of performance response to a single training bout". In: *J. Appl. Physiol.* 92.2, pp. 572–580 (cit. on pp. 64, 65, 68).

Busso, Thierry, Robin Candau, and Jean-René Lacour (1994). "Fatigue and fitness modelled from the effects of training on performance". In: *Eur. J. Appl. Physiol. Occup. Physiol.* 69.1, pp. 50–54 (cit. on pp. 64, 80).

Busso, Thierry, CLAUDE Carasso, and JEAN-RENE Lacour (1991). "Adequacy of a systems structure in the modeling of training effects on performance". In: *J. Appl. Physiol.* 71.5, pp. 2044–2049 (cit. on pp. 64, 69, 162).

Busso, Thierry, Christian Denis, Régis Bonnefoy, et al. (1997). "Modeling of adaptations to physical training by using a recursive least squares algorithm". In: *J. Appl. Physiol.* (cit. on pp. 61, 65, 66, 68, 69, 83).

Busso, Thierry, K Häkkinen, A Pakarinen, et al. (1990). "A systems model of training responses and its relationship to hormonal responses in elite weight-lifters". In: *Eur. J. Appl. Physiol. Occup. Physiol.* 61.1, pp. 48–54 (cit. on p. 64).

Busso, Thierry, K Häkkinen, A Pakarinen, et al. (1992). "Hormonal adaptations and modelled responses in elite weightlifters during 6 weeks of training". In: *Eur. J. Appl. Physiol. Occup. Physiol.* 64.4, pp. 381–386 (cit. on p. 64).

Busso, Thierry and L Thomas (2006). "Using mathematical modeling in training planning". In: *Int. J. Sports Physiol. Perform.* 1.4, pp. 400–405 (cit. on pp. 60, 61, 162).

Calvert, Thomas W, Eric W Banister, Margaret V Savage, et al. (1976). "A systems model of the effects of training on physical performance". In: *IEEE Trans. Syst. Man. Cybern.* 2, pp. 94–102 (cit. on pp. 32, 59, 60).

Candau, R, A Belli, GY Millet, et al. (1998). "Energy cost and running mechanics during a treadmill run to voluntary exhaustion in humans". In: *Eur. J Appl. Physiol. Occup. Physiol.* 77.6, pp. 479–485 (cit. on p. 197).

Cardinale, Marco and Matthew C Varley (2017). "Wearable training-monitoring technology: applications, challenges, and opportunities". In: *Ind. J. Sports Physiol. Perform.* 12.s2, S2–55 (cit. on p. 166).

Carrard, Apolline, Elisa Fontana, and Davide Malatesta (2018). "Mechanical Determinants of the U-Shaped Speed-Energy Cost of Running Relationship". In: *Front. Physiol.* 9, p. 1790 (cit. on pp. 197, 200).

Carrard, Justin, Petr Kloucek, and Boris Gojanovic (2020). "Modelling training adaptation in swimming using artificial neural network geometric optimisation". In: *Sports* 8.1, p. 8 (cit. on pp. 33, 85).

Casamichana, David, Julen Castellano, Julio Calleja-Gonzalez, et al. (2013). "Relationship between indicators of training load in soccer players". In: *J. Strength Cond. Res.* 27.2, pp. 369–374 (cit. on p. 159).

Cavagna, GA (1975). "Force platforms as ergometers". In: *J. Appl. Physiol.* 39.1, pp. 174–179 (cit. on pp. 192, 195).

Cavagna, GA and M Kaneko (1977). "Mechanical work and efficiency in level walking and running". In: *J. Physiol.* 268.2, pp. 467–481 (cit. on pp. 198, 203).

Cerqueira, Vitor, Luis Torgo, and Igor Mozetič (2020). "Evaluating time series forecasting models: An empirical study on performance estimation methods". In: *Mach. Learn.* 109.11, pp. 1997–2028 (cit. on p. 141).

Chai, Tianfeng and Roland R Draxler (2014). "Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature". In: *Geosci. Model Dev.* 7.3, pp. 1247–1250 (cit. on p. 63).

Chalencon, Sébastien, Vincent Pichot, Frédéric Roche, et al. (2015). "Modeling of performance and ANS activity for predicting future responses to training". In: *Eur. J. Appl. Physiol.* 115.3, pp. 589–596 (cit. on pp. 66, 75, 139, 162).

Chandler, RM, HK Byrne, JG Patterson, et al. (1994). "Dietary supplements affect the anabolic hormones after weight-training exercise". In: *J. Appl. Physiol.* 76.2, pp. 839–845 (cit. on p. 46).

Chelly, Souhaiel M and Christian Denis (2001). "Leg power and hopping stiffness: relationship with sprint running performance". In: *Med. Sci. Sports Exerc.* 33.2, pp. 326–333 (cit. on pp. 188, 265).

Chen, Trevor C and SANDY S HSIEH (2001). "Effects of a 7-day eccentric training period on muscle damage and inflammation". In: *Med. Sci. Sports Exerc.* 33.10, pp. 1732–1738 (cit. on p. 48).

Chiu, Loren ZF and Jacque L Barnes (2003). "The fitness-fatigue model revisited: Implications for planning short-and long-term training". In: *Strength Cond. J.* 25.6, pp. 42–51 (cit. on p. 64).

Chow, John W (1999). "Knee joint forces during isokinetic knee extensions: a case study". In: *Clin. Biomech.* 14.5, pp. 329–338 (cit. on p. 92).

Christ, Maximilian, Nils Braun, and Julius Neuffer (n.d.). "Overview on Time series feature extraction (tsfresh–a python package)". In: (). URL: `https://tsfresh.readthedocs.io/en/latest/text/list_of_features.html` (cit. on p. 184).

Christ, Maximilian, Nils Braun, Julius Neuffer, et al. (2018). "Time series feature extraction on basis of scalable hypothesis tests (tsfresh–a python package)". In: *Neurocomputing* 307, pp. 72–77 (cit. on p. 184).

Clarke, David C and Philip F Skiba (2013). "Rationale and resources for teaching the mathematical modeling of athletic training and performance". In: *Adv. Physiol. Educ.* (cit. on pp. 66, 72, 160).

Claudino, João Gustavo, Daniel de Oliveira Capanema, Thiago Vieira de Souza, et al. (2019). "Current approaches to the use of artificial intelligence for injury risk assessment and performance prediction in team sports: a systematic review". In: *Sports Med. Open* 5.1, pp. 1–12 (cit. on p. 166).

Coffey, Vernon G and John A Hawley (2007). "The molecular bases of training adaptation". In: *Sports Med.* 37.9, pp. 737–763 (cit. on p. 78).

Coggan, AR (2003). "Training and racing using a power meter: An introduction Level II Coaching Manual (pp. 123-145)". In: *Colorado Springs, CO: USA cycling* (cit. on pp. 44–46).

Colby, Marcus J, Brian Dawson, Jarryd Heasman, et al. (2014). "Accelerometer and GPS-derived running loads and injury risk in elite Australian footballers". In: *J. Strength Cond. Res.* 28.8, pp. 2244–2252 (cit. on p. 53).

Conceição, Filipe, Juvenal Fernandes, Martin Lewis, et al. (2016). "Movement velocity as a measure of exercise intensity in three lower limb exercises". In: *J. Sports Sci.* 34.12, pp. 1099–1106 (cit. on p. 51).

Cooper, DM, Colin Berry, Norman Lamarra, et al. (1985). "Kinetics of oxygen uptake and heart rate at onset of exercise in children". In: *J. Appl. Physiol.* 59.1, pp. 211–217 (cit. on p. 37).

Coppalle, Sullivan, Guillaume Rave, Abderraouf Ben Abderrahman, et al. (2019). "Relationship of pre-season training load with in-season biochemical markers, injuries and performance in professional soccer players". In: *Front. Physiol.* 10, p. 409 (cit. on p. 166).

Coutts, Aaron J and Rob Duffield (2010). "Validity and reliability of GPS devices for measuring movement demands of team sports". In: *J. Sci. Med. Sport* 13.1, pp. 133–135 (cit. on p. 52).

Cox, David Roxbee and Hilton David Miller (2017). *The theory of stochastic processes*. Routledge (cit. on p. 184).

Crang, Zachary L, Grant Duthie, Michael H Cole, et al. (2020). "The Validity and Reliability of Wearable Microtechnology for Intermittent Team Sports: A Systematic Review". In: *Sports Med.*, pp. 1–17 (cit. on p. 52).

Crewther, Blair, John Cronin, Justin Keogh, et al. (2008). "The salivary testosterone and cortisol response to three loading schemes". In: *J. Strength Cond. Res.* 22.1, pp. 250–255 (cit. on pp. 46, 47).

Cui, Zaixu and Gaolang Gong (2018). "The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features". In: *Neuroimage* 178, pp. 622–637 (cit. on pp. 83, 155).

Cunha, Felipe A, Adrian W Midgley, Lars R McNaughton, et al. (2016). "Effect of continuous and intermittent bouts of isocaloric cycling and running exercise on excess postexercise oxygen consumption". In: *J. Sci. Med. Sport* 19.2, pp. 187–192 (cit. on p. 42).

D'Antona, Giuseppe, Francesca Lanfranconi, Maria Antonietta Pellegrino, et al. (2006). "Skeletal muscle hypertrophy and structure and function of skeletal muscle fibres in male body builders". In: *J. Physiol.* 570.3, pp. 611–627 (cit. on p. 78).

D'Emanuele, Samuel, Nicola A Maffiuletti, Cantor Tarperi, et al. (2021). "Rate of Force Development as an Indicator of Neuromuscular Fatigue: A Scoping Review". In: *Front. Hum. Neurosci.* 15, p. 387 (cit. on pp. 96, 126).

Davis, James A, Ralph Rozenek, Derek M DeCicco, et al. (2007). "Comparison of three methods for detection of the lactate threshold". In: *Clin. Physiol. Funct. Imaging.* 27.6, pp. 381–384 (cit. on p. 38).

Davison, RC Richard, Ken A Van Someren, and Andrew M Jones (2009). "Physiological monitoring of the Olympic athlete". In: *J. Sports Sci.* 27.13, pp. 1433–1442 (cit. on p. 46).

Dayneka, Natalie L, Varun Garg, and William J Jusko (1993). "Comparison of four basic models of indirect pharmacodynamic responses". In: *J. Pharmacokinet. Biopharm.* 21.4, pp. 457–478 (cit. on p. 78).

De Salles, Belmiro Freitas, Roberto Simao, Fabrı+cio Miranda, et al. (2009). "Rest interval between sets in strength training". In: *Sports Med.* 39.9, pp. 765–777 (cit. on p. 50).

Dennis Jr, John E and Robert B Schnabel (1996). *Numerical methods for unconstrained optimization and nonlinear equations*. SIAM (cit. on p. 149).

Deschenes, Michael R and William J Kraemer (2002). "Performance and physiologic adaptations to resistance training". In: *Am. J. Phys. Med. Rehabil.* 81.11, S3–S16 (cit. on p. 88).

Di Blasio, Andrea, Pascal Izzicupo, Laura Tacconi, et al. (2014). "Acute and delayed effects of high intensity interval resistance training organization on cortisol and testosterone production". In: *J. Sports Med. Phys. Fitness.* 56.3, pp. 192–199 (cit. on p. 47).

Di Prampero, PE, G Atchou, J-C Brückner, et al. (1986). "The energetics of endurance running". In: *Eur. J. Appl. Physiol. and Occup. Physiol.* 55.3, pp. 259–266 (cit. on p. 196).

Doyon, Kristopher H, Stephane Perrey, Daijiro Abe, et al. (2001). "Field testing of in Cross-Country skiers with portable breath-by-breath system". In: *Can. J. Appl. Physiol.* 26.1, pp. 1–11 (cit. on pp. 94, 194).

Dufaux, B, G Assmann, H Schachten, et al. (1982). "The delayed effects of prolonged physical exercise and physical training on cholesterol level". In: *Eur. J. Appli. Physiol. Occup. Physiol.* 48.1, pp. 25–29 (cit. on p. 68).

Duffield, Rob, Brian Dawson, HC Pinnington, et al. (2004). "Accuracy and reliability of a Cosmed K4b2 portable gas analysis system". In: *J. Sci. Med. Sport* 7.1, pp. 11–22 (cit. on pp. 94, 194, 201).

Ebersole, Kyle T, Kristian M O'Connor, and Andrew P Wier (2006). "Mechanomyographic and electromyographic responses to repeated concentric muscle actions of the quadriceps femoris". In: *J. Electromyogr. Kinesiol.* 16.2, pp. 149–157 (cit. on p. 119).

Edelmann-Nusser, Jürgen, Andreas Hohmann, and Bernd Henneberg (2002). "Modeling and prediction of competitive performance in swimming upon neural networks". In: *Eur. J. Sport Sci.* 2.2, pp. 1–10 (cit. on pp. 33, 84).

Edwards, S (1993). "High performance training and racing". In: *The heart rate monitor book* 12.2, pp. 113–123 (cit. on p. 39).

Egan, Alison D, Jason B Winchester, Carl Foster, et al. (2006). "Using session RPE to monitor different methods of resistance exercise". In: *J. Sports Sci. Med.* 5.2, p. 289 (cit. on p. 133).

Egan, Brendan and Juleen R Zierath (2013). "Exercise metabolism and the molecular regulation of skeletal muscle adaptation". In: *Cell Metab.* 17.2, pp. 162–184 (cit. on p. 42).

Ehrmann, Fabian E, Craig S Duncan, Doungkamol Sindhusake, et al. (2016). "GPS and injury prevention in professional soccer". In: *J. Strength Cond. Res.* 30.2, pp. 360–367 (cit. on p. 53).

Eliakim, Eyal, Elia Morgulev, Ronnie Lidor, et al. (2020). "Estimation of injury costs: financial damage of English Premier League teams' underachievement due to injuries". In: *BMJ Open Sport Exerc. Med.* 6.1, e000675 (cit. on p. 32).

Engelen, Marielle, Janos Porszasz, Marshall Riley, et al. (1996). "Effects of hypoxic hypoxia on O2 uptake and heart rate kinetics during heavy exercise". In: *J. Appl. Physiol.* 81.6, pp. 2500–2508 (cit. on p. 37).

Enoka, ROGER M and Douglas G Stuart (1992). "Neurobiology of muscle fatigue". In: *J. Appl. Physiol.* 72.5, pp. 1631–1648 (cit. on p. 125).

Essig, David A, Nathan L Alderson, Michael A Ferguson, et al. (2000). "Delayed effects of exercise on the plasma leptin concentration". In: *Metabol. Clin. Exp.* 49.3, pp. 395–399 (cit. on p. 68).

Faigenbaum, Avery D and Gregory D Myer (2010). "Resistance training among young athletes: safety, efficacy and injury prevention effects". In: *Br. J. Sports Med.* 44.1, pp. 56–63 (cit. on p. 88).

Fanchini, Maurizio, Ivan Ferraresi, Roberto Modena, et al. (2016). "Use of the CR100 scale for session rating of perceived exertion in soccer and its interchangeability with the CR10". In: *Int. J. Sports Physiol. Perform.* 11.3, pp. 388–392 (cit. on p. 56).

Farinatti, Paulo TV and Antonio G Castinheiras Neto (2011). "The effect of between-set rest intervals on the oxygen uptake during and after resistance exercise sessions performed with large-and small-muscle mass". In: *J. Strength Cond. Res.* 25.11, pp. 3181–3190 (cit. on p. 106).

Farley, Claire T and Octavio Gonzalez (1996). "Leg stiffness and stride frequency in human running". In: *J. Biomech.* 29.2, pp. 181–186 (cit. on pp. 200, 207, 266).

Farrell, PETER A, JACK H Wilmore, EDWARD F Coyle, et al. (1979). "Plasma lactate accumulation and distance running performance". In: *Med. Sci. Sports* 11.4, pp. 338–44 (cit. on p. 38).

Feigenbaum, Matthew S and Michael L Pollock (1999). "Prescription of resistance training for health and disease." In: *Med. Sci. Sports Exerc.* 31.1, pp. 38–45 (cit. on p. 88).

Felser, S, M Behrens, S Fischer, et al. (2016). "Relationship between strength qualities and short track speed skating performance in young athletes". In: *Scand. J. Med. Sci. Sports* 26.2, pp. 165–171 (cit. on p. 145).

Fenn, Wallace O (1930). "Work against gravity and work due to velocity changes in running: Movements of the center of gravity within the body and foot pressure on the ground". In: *Am. J. Physiol.* 93.2, pp. 433–462 (cit. on p. 50).

Ferrari, Marco, Makii Muthalib, and Valentina Quaresima (2011). "The use of near-infrared spectroscopy in understanding skeletal muscle physiology: recent developments". In: *Philos. Trans. A Math. Phys. Eng. Sci.* 369.1955, pp. 4577–4590 (cit. on p. 96).

Fielding, ROGER A, MARIONA A Violan, LORI Svetkey, et al. (2000). "Effects of prior exercise on eccentric exercise-induced neutrophilia and enzyme release." In: *Med. Sci. Sports Exerc.* 32.2, pp. 359–364 (cit. on p. 48).

Fischer-Sonderegger, Karin, Wolfgang Taube, Martin Rumo, et al. (2021). "How far from the gold standard? Comparing the accuracy of a Local Position Measurement (LPM) system and a 15 Hz GPS to a laser for measuring acceleration and running speed during team sports". In: *PloS one* 16.4, e0250549 (cit. on p. 52).

Fitz-Clarke, John R, R Hugh Morton, and Eric W Banister (1991). "Optimizing athletic performance by influence curves". In: *J. Appl. Physiol.* 71.3, pp. 1151–1158 (cit. on p. 63).

Fleck, Steven (1988). "Cardiovascular adaptations to resistance training". In: *Med. Sci. Sports Exerc.* 20.5 (cit. on p. 88).

Forti, Louis Nuvagah, Evelien Van Roie, Rose Njemini, et al. (2017). "Effects of resistance training at different loads on inflammatory markers in young adults". In: *Eur. J. Appl. Physiol.* 117.3, pp. 511–519 (cit. on p. 48).

Foster, Carl, Jessica A Florhaug, Jodi Franklin, et al. (2001). "A new approach to monitoring exercise training". In: *J Strength Cond. Res.* 15.1, pp. 109–115 (cit. on pp. 56, 57).

Foster, Carl and Alejandro Lucia (2007). "Running economy". In: *Sports Med.* 37.4-5, pp. 316–319 (cit. on p. 207).

Foster, Carl, Jose A Rodriguez-Marroyo, and Jos J De Koning (2017). "Monitoring training loads: the past, the present, and the future". In: *Int. J. Sports Physiol. Perform.* 12.s2, S2–2 (cit. on p. 57).

Freund, H, S Oyono-Enguelle, A_ Heitz, et al. (1986). "Work rate-dependent lactate kinetics after exercise in humans". In: *J. Appl. Physiol.* 61.3, pp. 932–939 (cit. on p. 38).

Friedman, Jerome, Trevor Hastie, and Rob Tibshirani (2010). "Regularization paths for generalized linear models via coordinate descent". In: *J. Stat. Softw.* 33.1, p. 1 (cit. on p. 83).

Fry, Andrew C (2004). "The role of resistance exercise intensity on muscle fibre adaptations". In: *Sports Med.* 34.10, pp. 663–679 (cit. on pp. 49, 88).

Fry, Andrew C and William J Kraemer (1997). "Resistance exercise overtraining and overreaching". In: *Sports Med.* 23.2, pp. 106–129 (cit. on pp. 88, 162).

Fuller, Wayne A (2009). *Introduction to statistical time series*. Vol. 428. John Wiley & Sons (cit. on p. 184).

Gaesser, Glenn A and CEORGE A Brooks (1984). "Metabolic bases of excess post-exercise oxygen". In: *Med. Sci. Sports Exerc.* 16.1, pp. 29–43 (cit. on p. 42).

Gaesser, Glenn A and David C Poole (1996). "The slow component of oxygen uptake kinetics in humans". In: *Exerc. Sport Sci. Rev.* 24.1, pp. 35–70 (cit. on pp. 37, 198).

Ganter, Nico, Kerstin Witte, and Jürgen Edelmann-Nusser (2006). "Performance prediction in cycling using antagonistic models". In: *Int. J. Comput. Sci. Sport* 5.2, pp. 56–59 (cit. on p. 77).

Garcı+a-Pinillos, Felipe, Pedro Á Latorre-Román, Luis E Roche-Seruendo, et al. (2019). "Prediction of power output at different running velocities through the two-point method with the Stryd™ power meter". In: *Gait and posture* 68, pp. 238–243 (cit. on p. 192).

Garcı+a-Pinillos, FELIPE, LUIS E Roche-Seruendo, NEOL Marcén-Cinca, et al. (2018). "Absolute Reliability and Concurrent Validity of the Stryd System for the Assessment of Running Stride Kinematics at Different Velocities." In: *J. Strength Cond. Res.* (cit. on pp. 192, 204).

Garcı+a-Ramos, Amador, Belén Feriche, Carmen Calderón, et al. (2015). "Training load quantification in elite swimmers using a modified version of the training impulse method". In: *Eur. J. Sport Sci.* 15.2, pp. 85–93 (cit. on p. 39).

Garcı+a-Ramos, Amador, Francisco Luis Pestaña-Melero, Alejandro Pérez-Castilla, et al. (2018). "Differences in the load–velocity profile between 4 bench-press variants". In: *Int. J. Sports Physiol. Perf.* 13.3, pp. 326–331 (cit. on p. 51).

Gaudino, Paolo, F Marcello Iaia, Anthony J Strudwick, et al. (2015). "Factors influencing perception of effort (session rating of perceived exertion) during elite soccer training". In: *Int. J. Sports Physiol. Perform.* 10.7, pp. 860–864 (cit. on p. 52).

Genner, Kyle M and Matthew Weston (2014). "A comparison of workload quantification methods in relation to physiological responses to resistance exercise". In: *J. Strength Cond. Res.* 28.9, pp. 2621–2627 (cit. on pp. 49, 88, 103, 126).

Gillet, Nicolas, Sophie Berjot, Robert J Vallerand, et al. (2012). "Examining the motivation-performance relationship in competitive sport: a cluster-analytic approach." In: *Int. J. Sport Exerc. Psychol.* 43.2, p. 79 (cit. on p. 158).

Gillette, Cynthia A, Richard C Bullough, and Christopher L Melby (1994). "Postexercise energy expenditure in response to acute aerobic or resistive exercise". In: *Int. J Sport Nut.* 4.4, pp. 347–360 (cit. on p. 43).

Gómez-Carmona, Carlos D, Alejandro Bastida-Castillo, Sergio J Ibáñez, et al. (2020). "Accelerometry as a method for external workload monitoring in invasion team sports. A systematic review". In: *PloS one* 15.8, e0236643 (cit. on pp. 52, 53, 166).

Green, HJ, RL Hughson, GW Orr, et al. (1983). "Anaerobic threshold, blood lactate, and muscle metabolites in progressive exercise". In: *J. Appl. Physiol.* 54.4, pp. 1032–1038 (cit. on p. 39).

Greer, Beau Kjerulf, Prawee Sirithienthad, Robert J Moffatt, et al. (2015). "EPOC comparison between isocaloric bouts of steady-state aerobic, intermittent aerobic,

and resistance training". In: *Res. Q. Exerc. Sport.* 86.2, pp. 190–195 (cit. on p. 43).

Grgic, Jozo, Brad J Schoenfeld, Mislav Skrepnik, et al. (2018). "Effects of rest interval duration in resistance training on measures of muscular strength: a systematic review". In: *Sports Med.* 48.1, pp. 137–151 (cit. on p. 50).

Grömping, Ulrike (2009). "Variable importance assessment in regression: linear regression versus random forest". In: *Am. Stat.* 63.4, pp. 308–319 (cit. on p. 151).

Guizelini, Pedrode Camargo, Rafael Alves de Aguiar, Benedito Sérgio Denadai, et al. (2018). "Effect of resistance training on muscle strength and rate of force development in healthy older adults: a systematic review and meta-analysis". In: *Exp. Gerontol.* 102, pp. 51–58 (cit. on p. 112).

Haddad, Monoem, Georgios Stylianides, Leo Djaoui, et al. (2017). "Session-RPE method for training load monitoring: validity, ecological usefulness, and influencing factors". In: *Front. Neurosci.* 11, p. 612 (cit. on p. 57).

Hader, Karim, Michael C Rumpf, Maxime Hertzog, et al. (2019). "Monitoring the athlete match response: Can external load variables predict post-match acute and residual fatigue in soccer? A systematic review with meta-analysis". In: *Sports Med.-open* 5.1, pp. 1–19 (cit. on p. 188).

Haff, G Gregory (2010). "Quantifying workloads in resistance training: a brief review". In: *Strength Cond. J.* 10, pp. 31–40 (cit. on pp. 48, 50, 88).

Hakkinen, K and A Pakarinen (1993). "Acute hormonal responses to two different fatiguing heavy-resistance protocols in male athletes". In: *J. Appl. Physiol.* 74.2, pp. 882–887 (cit. on p. 47).

Häkkinen, K, M Alen, WJ Kraemer, et al. (2003). "Neuromuscular adaptations during concurrent strength and endurance training versus strength training". In: *Eur. J. Appl. Physiol.* 89.1, pp. 42–52 (cit. on pp. 88, 96, 126).

Häkkinen, K and A Pakarinen (1991). "Serum hormones in male strength athletes during intensive short term strength training". In: *Eur. J. Appl. Physiol.* 63.3, pp. 194–199 (cit. on p. 47).

Häkkinen, K, A Pakarinen, M Alen, et al. (1988). "Neuromuscular and hormonal responses in elite athletes to two successive strength training sessions in one day". In: *Eur. J. Appl. Physiol.* 57.2, pp. 133–139 (cit. on p. 47).

Häkkinen, K and Arto Pakarinen (1995). "Acute hormonal responses to heavy resistance exercise in men and women at different ages". In: *Int. J. Sports Med.* 16.08, pp. 507–513 (cit. on p. 46).

Häkkinen, Keijo and Paavo V Komi (1983). "Electromyographic and mechanical characteristics of human skeletal muscle during fatigue under voluntary and reflex conditions". In: *Electroencephalogr. Clin. Neurophysiol.* 55.4, pp. 436–444 (cit. on p. 123).

Halaki, Mark and Karen Ginn (2012). "Normalization of EMG signals: to normalize or not to normalize and what to normalize to". In: *Computational intelligence in electromyography analysis-a perspective on current applications and future challenges*. InTech Rijeka, pp. 175–194 (cit. on p. 99).

Hamaoka, Takafumi and Kevin K McCully (2019). "Review of early development of near-infrared spectroscopy and recent advancement of studies on muscle oxygenation and oxidative metabolism". In: *J Physiol. Sci.* 69.6, pp. 799–811 (cit. on pp. 96, 97).

Hamilton, Ryan J, Carl D Paton, and William G Hopkins (2006). "Effect of high-intensity resistance training on performance of competitive distance runners". In: *Int. J. Sports Physiol. Perform.* 1.1, pp. 40–49 (cit. on p. 88).

Hansen, S, T Kvorning, M Kjaer, et al. (2001). "The effect of short-term strength training on human skeletal muscle: the importance of physiologically elevated hormone levels". In: *Scand. J. Med. Sci. Sport* 11.6, pp. 347–354 (cit. on p. 46).

Harris, Chad, Mark Debeliso, and Kent J Adams (2003). "THE EFFECTS OF RUNNING SPEED ON THE METABOLIC AND MECHANICAL ENERGY COSTS OF RUNNING." In: *J. Exerc. Physiol. Online* 6.3 (cit. on p. 197).

Harris, RC, RHT Edwards, E Hultman, et al. (1976). "The time course of phosphorylcreatine resynthesis during recovery of the quadriceps muscle in man". In: *Pflugers Arch.* 367.2, pp. 137–142 (cit. on p. 50).

Hartigan, John A and Manchek A Wong (1979). "Algorithm AS 136: A k-means clustering algorithm". In: *J. R. Stat. Soc. Ser. C. Appl. Stat.* 28.1, pp. 100–108 (cit. on p. 181).

Heckel, Zoltán, Tamás Atlasz, Éva Tékus, et al. (2019). "Monitoring exercise-induced muscle damage indicators and myoelectric activity during two weeks of knee extensor exercise training in young and old men". In: *PloS one* 14.11, e0224866 (cit. on p. 48).

Helgerud, Jan, Kjetill Høydal, Eivind Wang, et al. (2007). "Aerobic high-intensity intervals improve V˙ O2max more than moderate training". In: *Med Sci Sports Exerc.* 39.4, pp. 665–671 (cit. on p. 162).

Hellard, Philippe, Marta Avalos, Christophe Hausswirth, et al. (2013). "Identifying optimal overload and taper in elite swimmers over time". In: *J. Sports Sci. Med.* 12.4, p. 668 (cit. on pp. 82, 83).

Hellard, Philippe, Marta Avalos, Lucien Lacoste, et al. (2006). "Assessing the limitations of the Banister model in monitoring training". In: *J. Sports Sci.* 24.05, pp. 509–520 (cit. on pp. 32, 61, 63, 65, 73, 74, 77, 82, 83, 154, 157).

Hellard, Philippe, Charlotte Scordia, Marta Avalos, et al. (2017). "Modelling of optimal training load patterns during the 11 weeks preceding major competition in elite swimmers". In: *Appl. Physiol. Nutr. Metab.* 42.10, pp. 1106–1117 (cit. on pp. 41, 82, 83).

Helms, Eric R, Kedric Kwan, Colby A Sousa, et al. (2020). "Methods for Regulating and Monitoring Resistance Training". In: *J. Hum. Kinet.* 74.1, pp. 23–42 (cit. on pp. 48, 57, 88).

Henson, Lindsey C, David C Poole, and Brian J Whipp (1989). "Fitness as a determinant of oxygen uptake response to constant-load exercise". In: *Eur. J. Appl. Physiol. Occup. Physiol.* 59.1, pp. 21–28 (cit. on p. 37).

Hermens, Hermie J, Bart Freriks, Catherine Disselhorst-Klug, et al. (2000). "Development of recommendations for SEMG sensors and sensor placement procedures". In: *J. Electromyogr. Kinesiol.* 10.5, pp. 361–374 (cit. on p. 99).

Hickson, RC, HA Bomze, and JO Hollozy (1978). "Faster adjustment of O2 uptake to the energy requirement of exercise in the trained state". In: *J. Appl. Physiol.* 44.6, pp. 877–881 (cit. on p. 38).

Hill, AV and Hartley Lupton (1923). "Muscular exercise, lactic acid, and the supply and utilization of oxygen". In: *Q. J. Med.* 62, pp. 135–171 (cit. on p. 41).

Hochreiter, Sepp and Jürgen Schmidhuber (1997). "Long short-term memory". In: *Neural Comput.* 9.8, pp. 1735–1780 (cit. on p. 173).

Hoerl, Arthur E and Robert W Kennard (1970). "Ridge regression: Biased estimation for nonorthogonal problems". In: *Technometrics* 12.1, pp. 55–67 (cit. on pp. 83, 173).

Hoffman, Jay R, JOOHEE Im, KENNETH W Rundell, et al. (2003). "Effect of muscle oxygenation during resistance exercise on anabolic hormone response". In: *Med. Sci. Sports Exerc.* 35.11, pp. 1929–1934 (cit. on pp. 46, 47).

Hooper, Sue L and Laurel Traeger Mackinnon (1995). "Monitoring overtraining in athletes". In: *Sports Med.* 20.5, pp. 321–327 (cit. on p. 58).

Hooper, Sue L, Laurel Traeger Mackinnon, ALF Howard, et al. (1995). "Markers for monitoring overtraining and recovery." In: *Med. Sci. Sports Exerc.* (cit. on p. 58).

Howe, Samuel T., Robert J. Aughey, William G. Hopkins, et al. (2017). "Quantifying important differences in athlete movement during collision-based team sports: Accelerometers outperform Global Positioning Systems". In: *2017 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL)*, pp. 1–4. DOI: 10.1109/ISISS.2017.7935655 (cit. on p. 52).

Hughson, RICHARD L, KENNETH H Weisiger, and GEORGE D Swanson (1987). "Blood lactate concentration increases as a continuous function in progressive exercise". In: *J. Appl. Physiol.* 62.5, pp. 1975–1981 (cit. on p. 38).

Hurley, BF, JAMES M Hagberg, WILLIAM K Allen, et al. (1984). "Effect of training on blood lactate levels during submaximal exercise". In: *J. Appl. Physiol.* 56.5, pp. 1260–1264 (cit. on p. 38).

Hyndman, Rob J and George Athanasopoulos (2018). *Forecasting: principles and practice*. OTexts (cit. on p. 141).

IJ, H (2018). "Statistics versus machine learning". In: *Nat. Meth.* 15.4, p. 233 (cit. on p. 84).

Ille, Anne, Ingrid Selin, Manh-Cuong Do, et al. (2013). "Attentional focus effects on sprint start performance as a function of skill level". In: *J. Sports Sci.* 31.15, pp. 1705–1712 (cit. on p. 158).

Imbach, Frank (2019). *RunningPerfApp: a Shiny web-application for runners*. https://fimbach.shinyapps.io/RunningPerfApp/ (cit. on p. 208).

– (2020). *sysmod : An R package for dose-response modelling in sports*. https://github.com/fimbach/sysmod (cit. on p. 152).

Imbach, Frank, Robin Candau, Romain Chailan, et al. (2020). "Validity of the Stryd power meter in measuring running parameters at submaximal speeds". In: *Sports* 8.7, p. 103 (cit. on p. 208).

Impellizzeri, Franco M, Ermanno Rampinini, and Samuele M Marcora (2005). "Physiological assessment of aerobic training in soccer". In: *J. Sports Sci.* 23.6, pp. 583–592 (cit. on p. 31).

Ito, AKIRA, PAAVO V Komi, BERTIL Sjödin, et al. (1983). "Mechanical efficiency of positive work in running at different speeds." In: *Med. Sci. Sports Exerc.* 15.4, pp. 299–308 (cit. on p. 197).

Jacks, Dean E, James Sowash, John Anning, et al. (2002). "Effect of exercise at three exercise intensities on salivary cortisol". In: *J. Strength Cond. Res.* 16.2, pp. 286–289 (cit. on p. 103).

James, C, P Sacco, MV Hurley, et al. (1994). "An evaluation of different protocols for measuring the force-velocity relationship of the human quadriceps muscles". In: *Eur. J. Appl. Physiol. Occup. Physiol.* 68.1, pp. 41–47 (cit. on pp. 90, 92).

James, Gareth, Daniela Witten, Trevor Hastie, et al. (2013). *An introduction to statistical learning*. Vol. 112. Springer (cit. on p. 141).

Jaric, S (2015). "Force-velocity relationship of muscles performing multi-joint maximum performance tasks". In: *Int. J. Sports Med.* 36.09, pp. 699–704 (cit. on p. 135).

Jeffreys, Harold (1961). *Theory of probability, Clarendon* (cit. on pp. 263, 264).

Jeffries, Annie C, Lee Wallace, Aaron J Coutts, et al. (2020). "Athlete-reported outcome measures for monitoring training responses: A systematic review of risk of bias and measurement property quality according to the COSMIN guidelines". In: *Int. J. Sports Physiol. Perform.* 15.9, pp. 1203–1215 (cit. on pp. 58, 59, 212, 215).

Jennings, Denise, Stuart Cormack, Aaron J Coutts, et al. (2010). "The validity and reliability of GPS units for measuring distance in team sport specific running patterns". In: *Int. J. Sports Pysiol. Perform.* 5.3, pp. 328–341 (cit. on pp. 52, 166).

Jones, Andrew M and David C Poole (2005). "Oxygen uptake dynamics: from muscle to mouth—an introduction to the symposium". In: *Med. Sci. Sports Exerc.* 37.9, pp. 1542–1550 (cit. on p. 37).

Jones, Marc R, Daniel J West, Blair T Crewther, et al. (2015). "Quantifying positional and temporal movement patterns in professional rugby union using global positioning system". In: *Eur. J. Sport Sci.* 15.6, pp. 488–496 (cit. on p. 52).

Jung, Won-Sang, Hyejung Hwang, Jisu Kim, et al. (2019). "Effect of interval exercise versus continuous exercise on excess post-exercise oxygen consumption during energy-homogenized exercise on a cycle ergometer". In: *J. Exerc. Nutrition Biochem.* 23.2, p. 45 (cit. on p. 42).

Kaiser, Henry F (1960). "The application of electronic computers to factor analysis". In: *Educ. Psychol. Meas.* 20.1, pp. 141–151 (cit. on p. 150).

Kallus, Konrad Wolfgang (1995). *Erholungs-Belastungs-Fragebogen: EBF*. Swets Test Services Frankfurt, Germany (cit. on p. 58).

Kallus, Konrad Wolfgang and Michael Kellmann (2016). *The recovery-stress questionnaires: user manual*. Pearson London, UK: (cit. on p. 58).

Kalman, Rudolph Emil (1960). "A new approach to linear filtering and prediction problems". In: *J. Basic Eng.* (cit. on p. 73).

Kaminsky, LA, S Padjen, and J LaHam-Saeger (1990). "Effect of split exercise sessions on excess post-exercise oxygen consumption." In: *Br. J. Sports Med.* 24.2, pp. 95–98 (cit. on p. 42).

Kang, Jie, Jay R Hoffman, Joohee Im, et al. (2005). "Evaluation of physiological responses during recovery following three resistance exercise programs". In: *J. Strength Cond. Res.* 19.2, pp. 305–309 (cit. on p. 106).

Karvonen, Juha and Timo Vuorimaa (1988). "Heart rate and exercise intensity during sports activities". In: *Sports Med.* 5.5, pp. 303–311 (cit. on pp. 37, 38).

Keir, Daniel A, Juan M Murias, Donald H Paterson, et al. (2014). "Breath-by-breath pulmonary O2 uptake kinetics: effect of data processing on confidence in estimating model parameters". In: *Exp. Physiol.* 99.11, pp. 1511–1522 (cit. on p. 194).

Kelley, Ken and Scott E Maxwell (2003). "Sample size for multiple regression: obtaining regression coefficients that are accurate, not simply significant." In: *Psychol. Methods* 8.3, p. 305 (cit. on pp. 83, 155).

Kibler, WB, TJ Chandler, and ES Stracener (1992). "Musculoskeletal adaptations and injuries due to overtraining." In: *Exerc. Sport Sci. Rev.* 20, pp. 99–126 (cit. on p. 88).

Kiens, Bente and Erik A Richter (1998). "Utilization of skeletal muscle triacylglycerol during postexercise recovery in humans". In: *Am. J. Physiol.* 275.2, E332–E337 (cit. on p. 42).

Knobbe, Arno, Jac Orie, Nico Hofman, et al. (2017). "Sports analytics for professional speed skating". In: *Data Min. Knowl. Discov.* 31.6, pp. 1872–1902 (cit. on pp. 82–84).

Knuttgen, HG, FB Petersen, and K Klausen (1971). "Exercise with concentric and eccentric muscle contractions". In: *Acta Paediatr. Scand. Suppl.* 60, pp. 42–46 (cit. on p. 37).

Kolossa, D, MA Bin Azhar, C Rasche, et al. (2017). "Performance estimation using the fitness-fatigue model with kalman filter feedback". In: *Int. J. Comput. Sci. Sport* (cit. on pp. 61, 73–75, 254, 255).

Kosmidis, Ioannis and Louis Passfield (2015). "Linking the performance of endurance runners to training and physiological effects via multi-resolution elastic net". In: *arXiv preprint arXiv:1506.01388* (cit. on p. 83).

Kouvaris, Kostas, Jeff Clune, Loizos Kounios, et al. (2017). "How evolution learns to generalise: Using the principles of learning theory to understand the evolution of developmental organisation". In: *PLoS Comput. Biol.* 13.4, e1005358 (cit. on p. 138).

Kraemer, William J (1987). "Endocrine responses to resistance exercise". In: (cit. on p. 46).

Kraemer, William J, Michael R Deschenes, and Steven J Fleck (1988). "Physiological adaptations to resistance exercise". In: *Sports Med.* 6.4, pp. 246–256 (cit. on pp. 88, 129).

Kraemer, WILLIAM J, Joseph E Dziados, LOUIS J Marchitelli, et al. (1993). "Effects of different heavy-resistance exercise protocols on plasma beta-endorphin concentrations". In: *J. Appl. Physiol.* 74.1, pp. 450–459 (cit. on p. 103).

Kraemer, William J, STEVEN J Fleck, Joseph E Dziados, et al. (1993). "Changes in hormonal concentrations after different heavy-resistance exercise protocols in women". In: *J. Appl. Physiol.* 75.2, pp. 594–604 (cit. on pp. 46, 47).

Kraemer, William J, SE Gordon, SJ Fleck, et al. (1991). "Endogenous anabolic hormonal and growth factor responses to heavy resistance exercise in males and females". In: *Int. J. Sports Med.* 12.02, pp. 228–235 (cit. on p. 46).

Kraemer, William J, Keijo Häkkinen, Robert U Newton, et al. (1998). "Acute hormonal responses to heavy resistance exercise in younger and older men". In: *Eur. J. Appl. Physiol. Occup. Physiol.* 77.3, pp. 206–211 (cit. on pp. 46, 47).

Kraemer, William J, Keijo Häkkinen, Robert U Newton, et al. (1999). "Effects of heavy-resistance training on hormonal response patterns in younger vs. older men". In: *J. Appl. Physiol.* 87.3, pp. 982–992 (cit. on p. 102).

Kraemer, William J, Louis Marchitelli, Scott E Gordon, et al. (1990). "Hormonal and growth factor responses to heavy resistance exercise protocols". In: *J. Appl. Physiol.* 69.4, pp. 1442–1450 (cit. on p. 46).

Kraemer, William J, John F Patton, Howard G Knuttgen, et al. (1989). "Hypothalamic-pituitary-adrenal responses to short-duration high-intensity cycle exercise". In: *J. Appl. Physiol.* 66.1, pp. 161–166 (cit. on p. 103).

Kraemer, William J and Nicholas A Ratamess (2005). "Hormonal responses and adaptations to resistance exercise and training". In: *Sports Med.* 35.4, pp. 339–361 (cit. on pp. 46, 47, 49, 50, 88).

Kraemer, WJ, AC Fry, BJ Warren, et al. (1992). "Acute hormonal responses in elite junior weightlifters". In: *Int. J. Sports Med.* 13.02, pp. 103–109 (cit. on p. 47).

Kraemer, WJ, BJ Noble, MJ Clark, et al. (1987). "Physiologic responses to heavy-resistance exercise with very short rest periods". In: *Int. J. Sports Med.* 8.04, pp. 247–252 (cit. on pp. 47, 102).

Kupperman, Natalie and Jay Hertel (2020). "Global positioning system–derived workload metrics and injury risk in team-based field sports: a systematic review". In: *J. Athl. Train.* 55.9, pp. 931–943 (cit. on p. 166).

Kuriki, Heloyse Uliam, Emanuelle Moraes Mello, Fabio Micolis De Azevedo, et al. (2012). *The relationship between electromyography and muscle force*. Citeseer (cit. on p. 120).

Kyröläinen, H, PV Komi, and A Belli (1995). "Mechanical efficiency in athletes during running". In: *Scand. J. Med. Sci. Sports* 5.4, pp. 200–208 (cit. on pp. 198, 203).

Lagally, Kristen M, Robert J Robertson, Kara I Gallagher, et al. (2002). "Perceived exertion, electromyography, and blood lactate during acute bouts of resistance exercise". In: *Med. Sci. Sports Exerc.* 34.3, pp. 552–559 (cit. on pp. 49, 133).

Lambert, EV, A St Clair Gibson, and TD Noakes (2005). "Complex systems model of fatigue: integrative homoeostatic control of peripheral physiological systems during exercise in humans". In: *Br. J. Sports Med.* 39.1, pp. 52–62 (cit. on pp. 32, 161).

Lambert, MI, ZH Mbambo, and A St Clair Gibson (1998). "Heart rate during training and competition for longdistance running". In: *J. Sports Sci.* 16.sup1, pp. 85–90 (cit. on p. 44).

Laplante, Mathieu and David M Sabatini (2012). "mTOR signaling in growth control and disease". In: *Cell* 149.2, pp. 274–293 (cit. on p. 78).

Lara, Frank, Lee Shearer, Mason Coppi, et al. (2018). "Reliability Of A Running Power Meter Between Trials Of Submaximal Running On Three Different Surfaces: 1838 Board# 99 May 31 330 PM-500 PM". In: *Med. Sci. Sports Exerc.* 50.5S, p. 436 (cit. on pp. 192, 201).

Lee, Michael D and Eric-Jan Wagenmakers (2014). *Bayesian cognitive modeling: A practical course*. Cambridge university press (cit. on p. 263).

Leite, RD, J Prestes, C Rosa, et al. (2011). "Acute effect of resistance training volume on hormonal responses in trained men". In: *J. Sports Med. Phys. Fitness.* 51.2, p. 322 (cit. on p. 47).

Lemaire, A, M Ripamonti, M Ritz, et al. (2014). "Agreement of three vs. eight isokinetic preset velocities to determine knee extensor torque-and power-velocity relationships". In: *Isokin. Exerc. Sci.* 22.1, pp. 1–7 (cit. on pp. 90, 92).

LeSuer, Dale A, James H McCormick, Jerry L Mayhew, et al. (1997). "The accuracy of prediction equations for estimating 1-RM performance in the bench press, squat, and deadlift". In: *J. Strength Cond. Res.* 11, pp. 211–213 (cit. on p. 49).

Lever, Jake, Martin Krzywinski, and Naomi Altman (2016). "Points of significance: model selection and overfitting". In: *Nat. Methods* 13.9, pp. 703–705 (cit. on p. 138).

Linari, Marco, Elisabetta Brunello, Massimo Reconditi, et al. (2015). "Force generation by skeletal muscle is controlled by mechanosensing in myosin filaments". In: *Nature* 528.7581, pp. 276–279 (cit. on p. 112).

Livesey, Geoffrey and Marines Elia (1988). "Estimation of energy expenditure, net carbohydrate utilization, and net fat oxidation and synthesis by indirect calorimetry: evaluation of errors with special reference to the detailed composition of fuels". In: *Am. J. Clin. Nutr.* 47.4, pp. 608–628 (cit. on p. 196).

Loyola-Gonzalez, Octavio (2019). "Black-box vs. white-box: Understanding their advantages and weaknesses from a practical point of view". In: *IEEE Access* 7, pp. 154096–154113 (cit. on p. 164).

Lucia, A, J Hoyos, A Carvajal, et al. (1999). "Heart rate response to professional road cycling: the Tour de France". In: *Int. J. Sports Med.* 20.03, pp. 167–172 (cit. on p. 40).

Lucia, Alejandro, Jesús Hoyos, Alfredo Santalla, et al. (2003). "Tour de France versus Vuelta a Espana: which is harder?" In: *Med. Sci. Sports Exerc.* 35.5, pp. 872–878 (cit. on p. 40).

Ludwig, Melanie, Alexander Asteroth, Christian Rasche, et al. (2019). "Including the Past: Performance Modeling Using a Preload Concept by Means of the Fitness-Fatigue Model". In: *Int. J. Comput. Sci. Sport* 18.1, pp. 115–134 (cit. on pp. 61, 65–67, 75, 139, 157).

Luenberger, David (1971). "An introduction to observers". In: *IEEE Trans Automat Contr.* 16.6, pp. 596–602 (cit. on p. 161).

Lupo, Corrado, Alexandru Nicolae Ungureanu, Riccardo Frati, et al. (2019). "Player session rating of perceived exertion: a more valid tool than coaches' ratings to monitor internal training load in elite youth female basketball". In: *Int. J Sports Physiol. Perform.* 15.4, pp. 548–553 (cit. on p. 58).

Ma'as, Moehammad Dzaky Fauzan, Ahmad Zahi Ulul Azmi, et al. (2017). "Real-time muscle fatigue monitoring based on median frequency of electromyography signal". In: *5th International Conference on Instrumentation, Control, and Automation (ICA)*. IEEE, pp. 135–139 (cit. on p. 99).

Madeleine, Pascal, Prem Bajaj, K Søgaard, et al. (2001). "Mechanomyography and electromyography force relationships during concentric, isometric and eccentric contractions". In: *J. Electromyogr. Kinesiol.* 11.2, pp. 113–121 (cit. on p. 120).

Maffiuletti, Nicola A, Per Aagaard, Anthony J Blazevich, et al. (2016). "Rate of force development: physiological and methodological considerations". In: *Eur. J. Appl. Physiol.* 116.6, pp. 1091–1116 (cit. on pp. 96, 113).

Malone, James J, Ric Lovell, Matthew C Varley, et al. (2017). "Unpacking the black box: applications and considerations for using GPS devices in sport". In: *Int. J. Sports Physiol. Perform.* 12.s2, S2–18 (cit. on p. 166).

Margaria, RODOLFO, HT Edwards, and DAVID BRUCE Dill (1933). "The possible mechanisms of contracting and paying the oxygen debt and the role of lactic acid in muscular contraction". In: *Am. J. Physiol.* 106.3, pp. 689–715 (cit. on p. 41).

Maritz, JS, JF Morrison, J Peter, et al. (1961). "A practical method of estimating an individual's maximal oxygen intake". In: *Ergonomics* 4.2, pp. 97–122 (cit. on p. 37).

Marquardt, Donald W and Ronald D Snee (1975). "Ridge regression in practice". In: *The American Statistician* 29.1, pp. 3–20 (cit. on p. 173).

Marston, Kieran J, Jeremiah J Peiffer, Michael J Newton, et al. (2017). "A comparison of traditional and novel metrics to quantify resistance training". In: *Sci. Rep.* 7.1, pp. 1–8 (cit. on pp. 50, 88, 102, 126, 133).

Matabuena, Marcos and Rosana Rodrı+guez-López (2019). "An improved version of the classical banister model to predict changes in physical condition". In: *Bull. Math. Biol.* 81.6, pp. 1867–1884 (cit. on pp. 69, 71).

Maupin, Danny, Ben Schram, Elisa Canetti, et al. (2020). "The relationship between acute: chronic workload ratios and injury risk in sports: a systematic review". In: *Open access J. Sports Med.* 11, p. 51 (cit. on p. 166).

Maybeck, Peter S (1990). "The Kalman filter: An introduction to concepts". In: *Autonomous robot vehicles*. Springer, pp. 194–204 (cit. on p. 73).

McBride, Jeffrey M, Grant O McCaulley, Prue Cormie, et al. (2009). "Comparison of methods to quantify volume during resistance exercise". In: *J. Strength Cond. Res.* 23.1, pp. 106–110 (cit. on p. 50).

McCaulley, Grant O, Jeffrey M McBride, Prue Cormie, et al. (2009). "Acute hormonal and neuromuscular responses to hypertrophy, strength and power type resistance exercise". In: *Eur. J. Appl. Physiol.* 105.5, pp. 695–704 (cit. on p. 47).

McCullagh, J and T Whitfort (2013). "An investigation into the application of Artificial Neural Networks to the prediction of injuries in sport". In: *Int. J. Sport Health Sci.* 7.7, pp. 356–360 (cit. on p. 85).

McCullagh, John et al. (2010). "Data mining in sport: A neural network approach". In: *Int. J. Sports Sci. Engineer.* 4.3, pp. 131–138 (cit. on p. 85).

McGuigan, Michael R, Alison D Egan, and Carl Foster (2004). "Salivary cortisol responses and perceived exertion during high intensity and low intensity bouts of resistance exercise". In: *J. Sports Sci. Med.* 3.1, p. 8 (cit. on p. 47).

McGuigan, Michael R and Jason B Winchester (2008). "The relationship between isometric and dynamic strength in college football players". In: *J. Sports Sci. Med.* 7.1, p. 101 (cit. on p. 112).

McGuigan, Michael R, Jason B Winchester, and Travis Erickson (2006). "The importance of isometric maximum strength in college wrestlers". In: *J. Sports Sci. Med.* 5.CSSI, p. 108 (cit. on p. 112).

McMahon, Bobby (2019). *Report Estimates The Cost Of Injuries To Premier League Players At $267M*. URL: `https://www.forbes.com/sites/bobbymcmahon/2019/08/22/report-shows-that-an-injury-to-a-premier-league-player-costs-on-average-350000/` (cit. on p. 32).

McMahon, TA and GC Cheng (1990). "The mechanics of running: how does stiffness couple with speed?" In: *J. Biomech.* 23, pp. 65–78 (cit. on pp. 192, 196, 198–200).

McNair, Douglas M, Maurice Lorr, Leo F Droppleman, et al. (1971). "Manual profile of mood states". In: (cit. on p. 58).

Méline, Thibaut, Ludovic Mathieu, Fabio Borrani, et al. (2019). "Systems model and individual simulations of training strategies in elite short-track speed skaters". In: *J. Sports Sci.* 37.3, pp. 347–355 (cit. on p. 160).

Miller, Benjamin F, Jens L Olesen, Mette Hansen, et al. (2005). "Coordinated collagen and muscle protein synthesis in human patella tendon and quadriceps muscle after exercise". In: *J. Physiol.* 567.3, pp. 1021–1033 (cit. on p. 73).

Millet, GP, RB Candau, B Barbier, et al. (2002). "Modelling the transfers of training effects on performance in elite triathletes". In: *Int. J. Sports Med.* 23.01, pp. 55–63 (cit. on p. 162).

Mitchell, Lachlan JG, Ben Rattray, John Fowlie, et al. (2020). "The impact of different training load quantification and modelling methodologies on performance predictions in elite swimmers". In: *Eur. J. Sport Sci.* 20.10, pp. 1329–1338 (cit. on pp. 33, 85, 139, 163).

Moore, Isabel S (2016). "Is there an economical running technique? A review of modifiable biomechanical factors affecting running economy". In: *Sports Med.* 46.6, pp. 793–807 (cit. on p. 207).

Morin, Jean-Benoit, Yann Le Mat, Cristian Osgnach, et al. (2021). "Individual acceleration-speed profile in-situ: A proof of concept in professional football players". In: *J. Biomech.* 123, p. 110524 (cit. on pp. 167, 170, 188).

Morin, Jean-Benoı+t, Georges Dalleau, Heikki Kyröläinen, et al. (2005). "A simple method for measuring stiffness during running". In: *J. Appl. Biomech.* 21.2, pp. 167–180 (cit. on p. 200).

Morin, Jean-Benoı+t, Pierre Samozino, K Zameziati, et al. (2007). "Effects of altered stride frequency and contact time on leg-spring behavior in human running". In: *J. Biomech.* 40.15, pp. 3341–3348 (cit. on p. 207).

Morton, R Hugh, John R Fitz-Clarke, and Eric W Banister (1990). "Modeling human performance in running". In: *J. Appl. Physiol.* 69.3, pp. 1171–1177 (cit. on pp. 62–64).

Mujika, Iñigo, Thierry Busso, Lucien Lacoste, et al. (1996). "Modeled responses to training and taper in competitive swimmers." In: *Med. Sci. Sports Exerc.* 28.2, pp. 251–258 (cit. on pp. 41, 64, 74, 154).

Mujika, Iñigo, Jean Claude Chatard, T Busso, et al. (n.d.). "Use of swim-training profiles and performance data to enhance training effectiveness". In: *J. Swim. Res.* 11 (), pp. 23–29 (cit. on p. 82).

MURAMATSU, Shigeru, Shuzo KATAO, and Ikuo HoMMA (1995). "Cardiorespiratory Responses and Mechanical Efficiency During Repeated Isokinetic Extension-Flexion Exercises of the Upper and Lower Limbs". In: *Showa Uni. J. Med. Sci.* 7.2, pp. 163–172 (cit. on p. 106).

Nagasawa, Takeshi, Takafumi Hamaoka, Takayuki Sako, et al. (2003). "A practical indicator of muscle oxidative capacity determined by recovery of muscle O 2 consumption using NIR spectroscopy". In: *Eur. J. Sport Sci.* 3.2, pp. 1–10 (cit. on p. 96).

Navalta, James W, Jeffrey Montes, Nathaniel G Bodell, et al. (2018). "Wearable Device Validity in Determining Step Count During Hiking and Trail Running". In: *J. Meas. Phys. Behav.* 1.2, pp. 86–93 (cit. on p. 192).

Navalta, James Wilfred, Jeffrey Montes, Nathaniel G Bodell, et al. (2019). "Reliability of Trail Walking and Running Tasks Using the Stryd Power Meter". In: *Int. J. Sports Med.* 40.08, pp. 498–502 (cit. on p. 197).

Nicolella, Daniel P, Lorena Torres-Ronda, Kase J Saylor, et al. (2018). "Validity and reliability of an accelerometer-based player tracking device". In: *PloS one* 13.2, e0191823 (cit. on p. 53).

Nilsson, Johnny, P Tesch, and Alf Thorstensson (1977). "Fatigue and EMG of repeated fast voluntary contractions in man". In: *Acta Physiol. Scand.* 101.2, pp. 194–198 (cit. on p. 123).

Noble, Bruce J, Gunnar A Borg, IRA Jacobs, et al. (1983). "A category-ratio perceived exertion scale: relationship to blood and muscle lactates and heart rate." In: *Med. Sci. Sports Exerc.* 15.6, pp. 523–528 (cit. on pp. 55, 159).

Nosaka, Kazunori and Mike Newton (2002). "Concentric or eccentric training effect on eccentric exercise-induced muscle damage". In: *Med. Sci. Sports Exerc.* 34.1, pp. 63–69 (cit. on p. 48).

Osgnach, Cristian, Stefano Poser, Riccardo Bernardini, et al. (2010). "Energy cost and metabolic power in elite soccer: a new match analysis approach". In: *Med. Sci. Sports Exerc.* 42.1, pp. 170–178 (cit. on pp. 52, 166).

Osgnach, Cristian and Pietro Enrico di Prampero (2018). "Metabolic power in team sports-Part 2: aerobic and anaerobic energy yields". In: *Int. J. Sports Med.* 39.08, pp. 588–595 (cit. on p. 166).

Peñailillo, Luis, A Blazevich, H Numazawa, et al. (2015). "Rate of force development as a measure of muscle damage". In: *Sand. J. Med. Sci. Sports* 25.3, pp. 417–427 (cit. on p. 96).

Pérez-Castilla, Alejandro, Antonio Piepoli, Gabriel Delgado-Garcı+a, et al. (2019). "Reliability and concurrent validity of seven commercially available devices for the assessment of movement velocity at different intensities during the bench press". In: *J. Strength Cond. Res.* 33.5, pp. 1258–1265 (cit. on p. 50).

Perl, Jürgen (2000). "Antagonistic adaptation systems: An example of how to improve understanding and simulating complex system behaviour by use of meta-models and on line-simulation". In: *16th IMACS Congress*. Citeseer (cit. on p. 76).

– (2001). "PerPot: A metamodel for simulation of load performance interaction". In: *Eur. J. Sport Sci.* 1.2, pp. 1–13 (cit. on pp. 75–77).

Perl, Jürgen, Markus Tilp, Arnold Baca, et al. (2013). "Neural networks for analysing sports games". In: *Routledge handbook of sports performance analysis*. Routledge, pp. 255–265 (cit. on p. 85).

Péronnet, François and Guy Thibault (1989). "Mathematical analysis of running performance and world running records". In: *J. Appl. Physiol.* 67.1, pp. 453–465 (cit. on pp. 209–211).

Perrey, S (2010). "Time course of oxygen uptake at exercise". In: *Exercise physiology: From a cellular to an integrative approach*. Vol. 75. IOS Press, pp. 285–298 (cit. on p. 201).

Perrey, Stephane and Marco Ferrari (2018). "Muscle oximetry in sports science: a systematic review". In: *Sports Med.* 48.3, pp. 597–616 (cit. on p. 96).

Pfeiffer, Mark (2008). "Modeling the relationship between training and performance-a comparison of two antagonistic concepts". In: *Int. J. Comput. Sci. Sport* 7.2, pp. 13–32 (cit. on pp. 64, 77).

Pfeiffer, Mark and Andreas Hohmann (2012). "Applications of neural networks in training science". In: *Hum. Mov. Sci.* 31.2, pp. 344–359 (cit. on p. 85).

Pfeiffer, Mark and Jürgen Perl (2006). "Analysis of tactical structures in team handball by means of artificial neural networks". In: *Int. J. Comput. Sci. Sport* 5.1, pp. 4–14 (cit. on p. 85).

Philippe, Antony G, Fabio Borrani, Anthony MJ Sanchez, et al. (2019). "Modelling performance and skeletal muscle adaptations with exponential growth functions during resistance training". In: *J. Sports Sci.* 37.3, pp. 254–261 (cit. on pp. 73, 88, 145, 155).

Phinyomark, A., Sirinee Thongpanja, Huosheng Hu, et al. (2012). "The Usefulness of Mean and Median Frequencies in Electromyography Analysis". In: (cit. on p. 99).

Piatrikova, Eva, Nicholas J Willsmer, Marco Altini, et al. (2021). "Monitoring the Heart Rate Variability Responses to Training Loads in Competitive Swimmers Using a Smartphone Application and the Banister Impulse-Response Model". In: *Int. J. Sports Physiol. Perform.* 16.6, pp. 787–795 (cit. on pp. 64, 75, 154).

Poole, David C, Thomas J Barstow, Glenn A Gaesser, et al. (1994). "VO2 slow component: physiological and functional significance." In: *Med. Sci. Sports Exerc.* 26.11, pp. 1354–1358 (cit. on p. 198).

Prampero, Pietro Enrico di and Cristian Osgnach (2018). "Metabolic power in team sports-part 1: an update". In: *Int. J. Sports Med.* 39.08, pp. 581–587 (cit. on p. 166).

Pullinen, TEEMU, ANTTI Mero, PIRKKO Huttunen, et al. (2002). "Resistance exercise-induced hormonal responses in men, women, and pubescent boys". In: *Med. Sci. Sports Exerc.* 34.5, pp. 806–813 (cit. on p. 46).

Pyne, David B, Tanya Boston, David T Martin, et al. (2000). "Evaluation of the Lactate Pro blood lactate analyser". In: *Eur. J. Appl. Physiol.* 82.1, pp. 112–116 (cit. on p. 95).

Racine, Jeff (2000). "Consistent cross-validatory model-selection for dependent data: hv-block cross-validation". In: *J. Econom.* 99.1, pp. 39–61 (cit. on pp. 142, 159).

Rago, Vincenzo, João Brito, Pedro Figueiredo, et al. (2020). "Methods to collect and interpret external training load using microtechnology incorporating GPS in professional football: a systematic review". In: *Res. Sports Med.* 28.3, pp. 437–458 (cit. on p. 52).

Rahimi, Rahman, Mohammad Ghaderi, Bahman Mirzaei, et al. (2010). "Acute IGF-1, cortisol and creatine kinase responses to very short rest intervals between sets during resistance exercise to failure in men". In: *World Appl. Sci. J.* 8.10, pp. 1287–93 (cit. on p. 50).

Rahimi, Rahman, Mohammad Qaderi, Hassan Faraji, et al. (2010). "Effects of very short rest periods on hormonal responses to resistance exercise in men". In: *J. Strength Cond. Res.* 24.7, pp. 1851–1859 (cit. on p. 50).

Rahmani, Abderrahmane, Pierre Samozino, Jean-Benoit Morin, et al. (2018). "A simple method for assessing upper-limb force–velocity profile in bench press". In: *Int. J. Sports Physiol.Perform.* 13.2, pp. 200–207 (cit. on p. 135).

Ratamess, Nicholas, Brent Alvar, T Evetoch, et al. (2009). "Progression models in resistance training for healthy adults [ACSM position stand]". In: *Med. Sci. Sports Exerc.* 41.3, pp. 687–708 (cit. on p. 96).

Ratamess, Nicholas A, William J Kraemer, Jeff S Volek, et al. (2005). "Androgen receptor content following heavy resistance exercise in men". In: *J. Steroid Biochem. Mol. Biol.* 93.1, pp. 35–42 (cit. on p. 47).

Ratamess, Nicholas A, Joseph G Rosenberg, Jie Kang, et al. (2014). "Acute oxygen uptake and resistance exercise performance using different rest interval lengths: The influence of maximal aerobic capacity and exercise sequence". In: *J.Strength Cond. Res.* 28.7, pp. 1875–1888 (cit. on p. 106).

Ravé, Guillaume, Urs Granacher, Daniel Boullosa, et al. (2020). "How to Use Global Positioning Systems (GPS) Data to Monitor Training Load in the "Real World" of Elite Soccer". In: *Front. Physiol.* 11 (cit. on pp. 52, 166).

Reynolds, Jeff M, Toryanno J Gordon, and Robert A Robergs (2006). "Prediction of one repetition maximum strength from multiple repetition maximum testing and anthropometry". In: *J. Strength Cond. Res.* 20.3, pp. 584–592 (cit. on pp. 92, 93).

Rico-González, Markel, Asier Los Arcos, Filipe M Clemente, et al. (2020). "Accuracy and reliability of local positioning systems for measuring sport movement patterns in stadium-scale: A systematic review". In: *Appl. Sci.* 10.17, p. 5994 (cit. on p. 52).

Rodrigues, Bernardo M, Estélio Dantas, Belmiro Freitas de Salles, et al. (2010). "Creatine kinase and lactate dehydrogenase responses after upper-body resistance exercise with different rest intervals". In: *J. Strength Cond. Res.* 24.6, pp. 1657–1662 (cit. on p. 48).

Rodrı+guez-Rosell, David, Fernando Pareja-Blanco, Per Aagaard, et al. (2018). "Physiological and methodological aspects of rate of force development assessment in human skeletal muscle". In: *Clin. Physiol. Funct. Imaging.* 38.5, pp. 743–762 (cit. on p. 96).

Rossi, Alessio, Luca Pappalardo, Paolo Cintia, et al. (2018). "Effective injury forecasting in soccer with GPS training data and machine learning". In: *PloS one* 13.7, e0201264 (cit. on p. 166).

Roston, Warren L, Brian J Whipp, James A Davis, et al. (1987). "Oxygen uptake kinetics and lactate concentration during exercise in humans". In: *American Review of Respiratory Disease* 135.5, pp. 1080–1084 (cit. on p. 37).

Ruddy, Joshua, Anthony Shield, Nirav Maniar, et al. (2018). "Predictive modeling of hamstring strain injuries in elite Australian footballers". In: *Med. Sci. Sports Exerc.* 50.5, pp. 906–914 (cit. on p. 85).

Rudin, Cynthia (2019). "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead". In: *Nature Mach Intell.* 1.5, pp. 206–215 (cit. on p. 164).

Rusko, HK, A Pulkkinen, S Saalasti, et al. (2003). "Pre-prediction of EPOC: a tool for monitoring fatigue accumulation during exercise". In: *Med. Sci. Sports Exerc.* 35.5 Suppl 1, S183 (cit. on p. 41).

Ryan, Samuel, Thomas Kempton, and Aaron J Coutts (2020). "Data reduction approaches to athlete monitoring in professional Australian football". In: *Int. J. Sports Physiol. Perform.* 1.aop, pp. 1–7 (cit. on p. 166).

Ryan, Terence E, W Michael Southern, Mary Ann Reynolds, et al. (2013). "A cross-validation of near-infrared spectroscopy measurements of skeletal muscle oxidative capacity with phosphorus magnetic resonance spectroscopy". In: *J. Appl. Physiol.* 115.12, pp. 1757–1766 (cit. on p. 96).

Sako, Takayuki, Takafumi Hamaoka, Hiroyuki Higuchi, et al. (2001). "Validity of NIR spectroscopy for quantitatively measuring muscle oxidative metabolic rate in exercise". In: *J. Appl. Physiol.* 90.1, pp. 338–344 (cit. on p. 96).

"Leave-One-Out Cross-Validation" (2010). In: *Encyclopedia of Machine Learning*. Ed. by Claude Sammut and Geoffrey I. Webb. Boston, MA: Springer US, pp. 600–601. ISBN: 978-0-387-30164-8. DOI: 10.1007/978-0-387-30164-8_469. URL: https://doi.org/10.1007/978-0-387-30164-8_469 (cit. on p. 139).

Sanchez, Anthony MJ, Olivier Galbès, Frédérique Fabre-Guery, et al. (2013). "Modelling training response in elite female gymnasts and optimal strategies of overload training and taper". In: *J. Sports Sci.* 31.14, pp. 1510–1519 (cit. on p. 160).

Schneider, Christoph, Florian Hanakam, Thimo Wiewelhove, et al. (2018). "Heart rate monitoring in team sports—a conceptual framework for contextualizing heart rate measures for training and recovery prescription". In: *Front. Physiol.* 9, p. 639 (cit. on p. 85).

Schneider, Donald A, Andrew N Wing, and Norman R Morris (2002). "Oxygen uptake and heart rate kinetics during heavy exercise: a comparison between arm cranking and leg cycling". In: *Eur. J. Appl. Physiol.* 88.1, pp. 100–106 (cit. on p. 37).

Scott, Brendan R, Grant M Duthie, Heidi R Thornton, et al. (2016). "Training monitoring for resistance exercise: theory and applications". In: *Sports Med.* 46.5, pp. 687–698 (cit. on p. 48).

Scott, Macfarlane TU, Tannath J Scott, and Vincent G Kelly (2016). "The validity and reliability of global positioning systems in team sport: a brief review". In: *J. Strength Cond. Res.* 30.5, pp. 1470–1490 (cit. on p. 52).

Scott, Tannath J, Cameron R Black, John Quinn, et al. (2013). "Validity and reliability of the session-RPE method for quantifying training in Australian football: a comparison of the CR10 and CR100 scales". In: *J. Strength Cond. Res.* 27.1, pp. 270–276 (cit. on pp. 56, 57).

Shaner, Aaron A, Jakob L Vingren, Disa L Hatfield, et al. (2014). "The acute hormonal response to free weight and machine weight resistance exercise". In: *J. Strength Cond. Res.* 28.4, pp. 1032–1040 (cit. on p. 46).

Sharma, Amarnath, William F Ebling, and William J Jusko (1998). "Precursor-dependent indirect pharmacodynamic response model for tolerance and rebound phenomena". In: *J. Pharm. Sci.* 87.12, pp. 1577–1584 (cit. on p. 78).

Sharma, Amarnath and William J Jusko (1998). "Characteristics of indirect pharmacodynamic models and applications to clinical drug responses". In: *Br. J. Clin. Pharmacol.* 45.3, pp. 229–239 (cit. on p. 78).

Silva, António José, Aldo Manuel Costa, Paulo Moura Oliveira, et al. (2007). "The use of neural network technology to model swimming performance". In: *J. Sports Sci. Med.* 6.1, p. 117 (cit. on p. 85).

Singh, Favil, Carl Foster, David Tod, et al. (2007). "Monitoring different types of resistance training using session rating of perceived exertion". In: *Int. J. Sports Physiol. Perform.* 2.1, pp. 34–45 (cit. on p. 57).

Slawinski, Jean, Richard Heubert, Jacques Quievre, et al. (2008). "Changes in spring-mass model parameters and energy cost during track running to exhaustion". In: *J. Strength Cond. Res.* 22.3, pp. 930–936 (cit. on p. 207).

Smilios, Ilias, Theophilos Pilianidis, Michalis Karamouzis, et al. (2003). "Hormonal responses after various resistance exercise protocols". In: *Med. Sci. Sports Exerc.* 35.4, pp. 644–654 (cit. on pp. 46, 47).

Smith, Douglas B, Terry J Housh, Glen O Johnson, et al. (1998). "Mechanomyographic and electromyographic responses to eccentric and concentric isokinetic muscle actions of the biceps brachii". In: *Muscle Nerve.* 21.11, pp. 1438–1444 (cit. on p. 120).

Smith, Steven (2013). *Digital signal processing: a practical guide for engineers and scientists*. Elsevier (cit. on p. 194).

Snyder, Kristine L, Wyatt P Mohrman, James A Williamson, et al. (2018). "Methodological Flaws in Aubry, RL, Power, GA, and Burr, JF. An Assessment of Running Power as a Training Metric for Elite and Recreational Runners. J Strength Cond Res 32: 2258–2264, 2018". In: *J. Strength Cond. Res.* 32.12, e61 (cit. on p. 201).

Song, Zhe, Daniel R Moore, Nathan Hodson, et al. (2017). "Resistance exercise initiates mechanistic target of rapamycin (mTOR) translocation and protein complex co-localisation in human skeletal muscle". In: *Sci. Rep.* 7.1, pp. 1–14 (cit. on p. 78).

Sontag, Eduardo D (2013). *Mathematical control theory: deterministic finite dimensional systems*. Vol. 6. Springer Science & Business Media (cit. on p. 161).

Spiering, Barry A, William J Kraemer, Jeffrey M Anderson, et al. (2008). "Resistance exercise biology". In: *Sports Med.* 38.7, pp. 527–540 (cit. on p. 46).

Stagno, Karl M, Rhys Thatcher, and Ken A Van Someren (2007). "A modified TRIMP to quantify the in-season training load of team sport players". In: *J. Sports Sci.* 25.6, pp. 629–634 (cit. on p. 40).

Staunton, Craig A, Grant Abt, Dan Weaving, et al. (2021). "Misuse of the term 'load'in sport and exercise science". In: *J. Sci. Med. Sport* (cit. on p. 31).

Stehle, Robert, Johannes Solzin, Bogdan Iorga, et al. (2009). "Insights into the kinetics of Ca 2+-regulated contraction and relaxation from myofibril studies". In: *Pflugers Archiv.* 458.2, pp. 337–357 (cit. on p. 115).

Stephens Hemingway, Benedict H, Katherine E Burgess, Eyad Elyan, et al. (2020). "The effects of measurement error and testing frequency on the fitness-fatigue model applied to resistance training: a simulation approach". In: *Int. J. Sports Sci. Coach.* 15.1, pp. 60–71 (cit. on p. 139).

Stone, Michael H, Meg Stone, and William A Sands (2007). *Principles and practice of resistance training.* Human Kinetics (cit. on pp. 64, 154).

Stryd (2017). *How to Lead the Pack: Running Power Meters and Quality Data* (cit. on pp. 201, 203, 204).

Sweet, Travis W, Carl Foster, Michael R McGuigan, et al. (2004). "Quantitation of resistance training using the session rating of perceived exertion method". In: *J. Strength Cond. Res.* 18.4, pp. 796–802 (cit. on p. 57).

Taha, Tim and Scott G Thomas (2003). "Systems modelling of the relationship between training and performance". In: *Sports Med.* 33.14, pp. 1061–1073 (cit. on p. 64).

Taylor, Sean J and Benjamin Letham (2018). "Forecasting at scale". In: *Am. Stat.* 72.1, pp. 37–45 (cit. on p. 176).

Tesi, C, F Colomo, S Nencini, et al. (2000). "The effect of inorganic phosphate on force generation in single myofibrils from rabbit skeletal muscle". In: *Biophys. J.* 78.6, pp. 3081–3092 (cit. on p. 115).

Theodoropoulos, John S, Jeremy Bettle, and Jonathan D Kosy (2020). "The use of GPS and inertial devices for player monitoring in team sports: A review of current and future applications". In: *Orthop. Rev.* 12.1 (cit. on p. 166).

Thomas, Luc and THIERRY Busso (2005). "A theoretical study of taper characteristics to optimize performance". In: *Med. Sci. Sports Exerc.* 37.9, p. 1615 (cit. on p. 160).

Thomas, Luc, Iñigo Mujika, and Thierry Busso (2008). "A model study of optimal training reduction during pre-event taper in elite swimmers". In: *J. Sports Sci.* 26.6, pp. 643–652 (cit. on pp. 41, 160).

– (2009). "Computer simulations assessing the potential performance benefit of a final increase in training during pre-event taper". In: *J. Strength Cond. Res.* 23.6, pp. 1729–1736 (cit. on p. 160).

Thongpanja, Sirinee, Angkoon Phinyomark, Pornchai Phukpattaranont, et al. (2013). "Mean and median frequency of EMG signal to determine muscle force based on time-dependent power spectrum". In: *Elektron. Elektrotech.* 19.3, pp. 51–56 (cit. on p. 99).

Thornton, M Kathleen and Jeffery A Potteiger (2002). "Effects of resistance exercise bouts of different intensities but equal work on EPOC". In: *Med. Sci. Sports Exerc.* 34.4, pp. 715–722 (cit. on p. 43).

Thorstensson, Alf, Gunnar Grimby, and Jan Karlsson (1976). "Force-velocity relations and fiber composition in human knee extensor muscles". In: *J. Appli. Physiol.* 40.1, pp. 12–16 (cit. on p. 123).

Thorstensson, Alf and Jan Karlsson (1976). "Fatiguability and fibre composition of human skeletal muscle". In: *Acta Physiol. Scand.* 98.3, pp. 318–322 (cit. on p. 123).

Tiao, George C and George EP Box (1981). "Modeling multiple time series with applications". In: *J. Am. Stat. Assoc.* 76.376, pp. 802–816 (cit. on p. 176).

Tibshirani, Robert (1996). "Regression shrinkage and selection via the lasso". In: *J. R. Stat. Soc. Series B Stat. Methodol.* 58.1, pp. 267–288 (cit. on p. 83).

Tiidus, PETER M and C DAVID Ianuzzo (1983). "Effects of intensity and duration of muscular exercise on delayed soreness and serum enzyme activities." In: *Med. Sci. Sports Exerc.* 15.6, pp. 461–465 (cit. on p. 48).

Torrents, Carlota, Natàlia Balagué, Jürgen Perl, et al. (2007). "Linear and nonlinear analysis of the traditional and differential strength training". In: *Balt. J. Sport Health Sci.* 3.66 (cit. on p. 77).

Tran, Quan T, David Docherty, and David Behm (2006). "The effects of varying time under tension and volume load on acute neuromuscular responses". In: *Eur. J. Appl. Physiol.* 98.4, pp. 402–410 (cit. on p. 51).

Tschakert, Gerhard and Peter Hofmann (2013). "High-intensity intermittent exercise: methodological and physiological aspects". In: *Int. J. Sports Physiol. Perform.* 8.6, pp. 600–610 (cit. on p. 41).

Tyska, MJ, DE Dupuis, WH Guilford, et al. (1999). "Two heads of myosin are better than one for generating force and motion". In: *Proc. Natl. Acad. Sci.* 96.8, pp. 4402–4407 (cit. on p. 112).

Vallance, Emmanuel, Nicolas Sutton-Charani, Abdelhak Imoussaten, et al. (2020). "Combining internal-and external-training-loads to predict non-contact injuries in soccer". In: *Appl. Sci.* 10.15, p. 5261 (cit. on p. 166).

Van Dijk, Hans and Ron Van Megen (2017). *The secret of running: Maximum performance gains through effective power metering and training analysis*. Meyer & Meyer Sport (cit. on pp. 44, 46).

Vanhelder, WoP, MW Radomski, and RC Goode (1984). "Growth hormone responses during intermittent weight lifting exercise in men". In: *Europ. J. Appl. Physiol.* 53.1, pp. 31–34 (cit. on p. 49).

Vanrenterghem, Jos, Niels Jensby Nedergaard, Mark A Robinson, et al. (2017). "Training load monitoring in team sports: a novel framework separating physiological and biomechanical load-adaptation pathways". In: *Sports Med.* 47.11, pp. 2135–2142 (cit. on p. 53).

Vanwanseele, Benedicte, Tim Op De Beéck, Kurt Schütte, et al. (2020). "Accelerometer Based Data Can Provide a Better Estimate of Cumulative Load During Running Compared to GPS Based Parameters". In: *Front. Sports Act. Living* 2 (cit. on p. 52).

Varley, Matthew C, Ian H Fairweather, and Robert J Aughey1 2 (2012). "Validity and reliability of GPS for measuring instantaneous velocity during acceleration, deceleration, and constant motion". In: *J. Sports Sci.* 30.2, pp. 121–127 (cit. on p. 52).

Varley, Matthew C, Arne Jaspers, Werner F Helsen, et al. (2017). "Methodological considerations when quantifying high-intensity efforts in team sport using global positioning system technology". In: *Int. J. Sports Physiol. Perform.* 12.8, pp. 1059–1068 (cit. on p. 52).

Vaswani, Ashish, Noam Shazeer, Niki Parmar, et al. (2017). "Attention is all you need". In: *Advances in neural information processing systems*, pp. 5998–6008 (cit. on p. 177).

Vehtari, Aki, Andrew Gelman, Daniel Simpson, et al. (2019). "Rank-normalization, folding, and localization: An improved ˆ R for assessing convergence of MCMC". In: *arXiv preprint arXiv:1903.08008*, p. 18 (cit. on p. 263).

Verbeke, Geert (1997). "Linear mixed models for longitudinal data". In: *Linear mixed models in practice*. Springer, pp. 63–153 (cit. on p. 82).

Verma, SK, LS Sidhu, and DK Kansal (1979). "A study of maximum oxygen uptake and heart rate during work and recovery as measured on cycle ergometer on national Indian sportsmen." In: *Br. J. Sports Med.* 13.1, pp. 24–28 (cit. on p. 37).

Vermeire, Kobe M, Freek Van de Casteele, Maxim Gosseries, et al. (2021). "The influence of different training load quantification methods on the fitness-fatigue model". In: *Int. J. Sports Physiol. Perform.* 1.aop, pp. 1–9 (cit. on pp. 64, 162).

Vingren, Jakob L, William J Kraemer, Nicholas A Ratamess, et al. (2010). "Testosterone physiology in resistance exercise and training". In: *Sports Med.* 40.12, pp. 1037–1053 (cit. on p. 46).

Viru, A (1984). "The mechanism of training effects: a hypothesis". In: *Int. J. Sports Med.* 5.05, pp. 219–227 (cit. on p. 53).

Viru, AA and M Viru (2001). "Biochemical monitoring of sport training: Human Kinetics". In: *Champaign, IL* (cit. on pp. 68, 69, 73).

Volek, Jeff S, William J Kraemer, Jill A Bush, et al. (1997). "Testosterone and cortisol in relationship to dietary nutrients and resistance exercise". In: *J. Appl. Physiol.* (cit. on p. 46).

Walker, Simon, Keijo Häkkinen, Robert U Newton, et al. (2020). "Acute responses of comprehensive gonadosteroids and corticosteroids to resistance exercise before and after 10 weeks of supervised strength training". In: *Exp. Physiol.* 105.3, pp. 438–448 (cit. on p. 88).

Wallace, Lee K, Katie M Slattery, and Aaron J Coutts (2009). "The ecological validity and application of the session-RPE method for quantifying training loads in swimming". In: *J. Strength Cond. Res.* 23.1, pp. 33–38 (cit. on p. 31).

Wallace, Lee K, Katie M Slattery, Franco M Impellizzeri, et al. (2014). "Establishing the criterion validity and reliability of common methods for quantifying training load". In: *J. Strength Cond. Res.* 28.8, pp. 2330–2337 (cit. on p. 57).

Wasserman, KARLMAN, WILLIAM L Beaver, and BRIAN J Whipp (1986). "Mechanisms and patterns of blood lactate increase during exercise in man." In: *Med. Sci. Sports Exerc.* 18.3, pp. 344–352 (cit. on p. 38).

Watkins, Casey M, Saldiam R Barillas, Megan A Wong, et al. (2017). "Determination of vertical jump as a measure of neuromuscular readiness and fatigue". In: *J. Strength Cond. Res.* 31.12, pp. 3305–3310 (cit. on p. 155).

Weakley, Jonathon, Bryan Mann, Harry Banyard, et al. (2021). "Velocity-based training: From theory to application". In: *Stren. Cond. J.* 43.2, pp. 31–49 (cit. on p. 51).

Weakley, Jonathon, Matthew Morrison, Amador Garcı+a-Ramos, et al. (2021). "The Validity and Reliability of Commercially Available Resistance Training Monitoring Devices: A Systematic Review". In: *Sports Med.*, pp. 1–60 (cit. on p. 50).

Welch, Greg, Gary Bishop, et al. (1995). "An introduction to the Kalman filter". In: (cit. on p. 73).

Willems, PA, GA Cavagna, and NC Heglund (1995). "External, internal and total work in human locomotion." In: *J. Exp. Biol.* 198.2, pp. 379–393 (cit. on pp. 197, 198, 203).

Williams, Mark A, William L Haskell, Philip A Ades, et al. (2007). "Resistance exercise in individuals with and without cardiovascular disease: 2007 update: a scientific statement from the American Heart Association Council on Clinical Cardiology and Council on Nutrition, Physical Activity, and Metabolism". In: *Circulation* 116.5, pp. 572–584 (cit. on p. 88).

Williams, Sean, Stephen West, Dan Howells, et al. (2018). "Modelling the HRV response to training loads in elite rugby sevens players". In: *J Sports Sci Med.* 17.3, p. 402 (cit. on p. 162).

Willmott, Cort J and Kenji Matsuura (2005). "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance". In: *Clim. Res.* 30.1, pp. 79–82 (cit. on p. 63).

Willoughby, Darryn S (2004). "Effects of heavy resistance training on myostatin mRNA and protein expression". In: *Med. Sci. Sports Exerc.* 36.4, pp. 574–582 (cit. on p. 47).

Wilmore, Jack H, Richard B Parr, Paul Ward, et al. (1978). "Energy cost of circuit weight training." In: *Med. Sci. Sports* 10.2, pp. 75–78 (cit. on p. 95).

Wisdom, Katrina M, Scott L Delp, and Ellen Kuhl (2015). "Use it or lose it: multiscale skeletal muscle adaptation to mechanical stimuli". In: *Biomech. Model Mechanobiol.* 14.2, pp. 195–215 (cit. on p. 53).

Witten, Daniela M and Robert Tibshirani (2009). "Covariance-regularized regression and classification for high dimensional problems". In: *J. R. Stat. Soc. Series B Stat. Methodol.* 71.3, pp. 615–636 (cit. on p. 83).

Witten, Ian H and Eibe Frank (2002). "Data mining: practical machine learning tools and techniques with Java implementations". In: *Acm Sigmod Record* 31.1, pp. 76–77 (cit. on p. 164).

Wolpert, Daniel M and Zoubin Ghahramani (2000). "Computational principles of movement neuroscience". In: *Nat. Neurosci.* 3.11, pp. 1212–1217 (cit. on p. 73).

Wolpert, David H (1992). "Stacked generalization". In: *Neural Netw.* 5.2, pp. 241–259 (cit. on p. 163).

– (2021). "What is important about the No Free Lunch theorems?" In: *Black Box Optimization, Machine Learning, and No-Free Lunch Theorems*. Springer, pp. 373–388 (cit. on p. 85).

Wolpert, David H and William G Macready (1997). "No free lunch theorems for optimization". In: *IEEE transactions on evolutionary computation* 1.1, pp. 67–82 (cit. on pp. 34, 85, 86).

Wood, Rachel Elise, Scott Hayter, David Rowbottom, et al. (2005). "Applying a mathematical model to training adaptation in a distance runner". In: *Eur. J. Appl. Physiol.* 94.3, pp. 310–316 (cit. on p. 66).

Zacks, RM (1973). "The mechanical efficiencies of running and bicycling against a horizontal impeding force". In: *Int. Z. Angew. Physiol.* 31.4, pp. 249–258 (cit. on p. 203).

Zafeiridis, Andreas, Ilias Smilios, Robert V Considine, et al. (2003). "Serum leptin responses after acute resistance exercise protocols". In: *J. Appl. Physiol.* 94.2, pp. 591–597 (cit. on p. 47).

Zihajehzadeh, Shaghayegh, Darrell Loh, Tien Jung Lee, et al. (2015). "A cascaded Kalman filter-based GPS/MEMS-IMU integration for sports applications". In: *Measurement* 73, pp. 200–210 (cit. on p. 52).

Zoladz, Jerzy A, Krzysztof Duda, and Joanna Majerczak (1998). "Oxygen uptake does not increase linearly at high power outputs during incremental exercise test in humans". In: *Eur. J. Appl. Physiol. Occup. Physiol.* 77.5, pp. 445–451 (cit. on p. 37).

Zou, Hui and Trevor Hastie (2005). "Regularization and variable selection via the elastic net". In: *J. R. Stat. Soc. Series B Stat. Methodol.* 67.2, pp. 301–320 (cit. on pp. 83, 150).

Zuccarelli, Lucrezia, Simone Porcelli, Letizia Rasica, et al. (2018). "Comparison between Slow Components of HR and VO2 Kinetics: Functional Significance." In: *Med. Sci. Sports Exerc.* 50.8, pp. 1649–1657 (cit. on p. 37).

ANNEXES

A. Appendix: Derivation of Equation 1.13

$$
\begin{aligned}
g'(t) &= w(t) - \frac{1}{\tau_g} g(t) \\
w(t) &= g'(t) + \frac{1}{\tau_g} g(t) \\
0 &= Cae^{ax} + \frac{1}{\tau_g} Ce^{ax} \\
0 &= Ce^{ax}(a + \frac{1}{\tau_g}) \\
g &= Ce^{-\frac{1}{\tau_g} x} \\
g(t) &= C(t)e^{\frac{-t}{\tau_g}} \\
w(t) &= g'(t) + \frac{1}{\tau_g} g(t) \\
w(t) &= C'e^{\frac{-t}{\tau_g}} \\
C' &= w(t)e^{\frac{t}{\tau_g}} \\
C(t) &= \int_0^t w(t)e^{\frac{s}{\tau_g}} ds \\
g(t) &= C(t)e^{\frac{-t}{\tau_g}} \\
g(t) &= e^{\frac{-t}{\tau_g}} \int_0^t w(s)e^{\frac{s}{\tau_g}} ds \\
g(t) &= \int_0^t w(s)e^{\frac{-1(t-s)}{\tau_g}} ds
\end{aligned}
$$

B. Derivation of Equations 1.33 to 1.35

$$
\begin{aligned}
g(t)e^{-\frac{1}{\tau_1}(t-k)} - g(k) &= \int_k^t \left[w(s) - \frac{1}{\tau_2}g(s-1)\right] e^{-\frac{1}{\tau_1}(s-k)} ds \\
g(t)e^{-\frac{1}{\tau_1}(t-k)} &= g(k) + \int_k^t \dots \\
g(t) &= g(k)e^{-\frac{1}{\tau_1}(t-k)} + e^{-\frac{1}{\tau_1}(t-k)} \int_k^t \dots \\
g(t) &= g(k)e^{-\frac{1}{\tau_1}(t-k)} + \int_k^t \left[w(s) - \frac{1}{\tau_2}g(s-1)\right] e^{-\frac{1}{\tau_1}(t-k-s+k)} ds \\
g(t) &= \dots \int_k^t \left[w(s) - \frac{1}{\tau_2}g(s-1)\right] e^{-\frac{1}{\tau_1}(t-s)} ds \\
g(k+1) &= g(k)e^{-\frac{1}{\tau_1}(k+1-k)} + \int_k^{k+1} \left[w(i) - \frac{1}{\tau_2}g(i-1)\right] e^{-\frac{1}{\tau_1}(k+1-i)} ds \\
g(k+1) &= g(k)e^{-\frac{1}{\tau_1}} + \sum_{i=k}^{k+1-i} \left[w(i) - \frac{1}{\tau_2}g(i-1)\right] e^{-\frac{1}{\tau_1}(k+1-i)} \\
g(k+1) &= g(k)e^{-\frac{1}{\tau_1}} + \left[w(k) - \frac{1}{\tau_2}g(k-1)\right] e^{-\frac{1}{\tau_1}(k+1-i)} \\
g(k+1) &= \dots e^{-\frac{1}{\tau_1}} \\
g(k+1) &= \left(g(k) + w(k) - \frac{1}{\tau_2}g(k-1)\right) e^{-\frac{1}{\tau_1}}
\end{aligned}
$$

$$
\begin{aligned}
g(t)e^{\frac{1t}{\tau_1}} - g(0) &= \int_0^t \left[w(s) - \frac{1}{\tau_2} g(s-1) \right] e^{\frac{1s}{\tau_1}} ds \\
g(t) &= \int_0^t \left[w(s) - \frac{1}{\tau_2} g(s-1) \right] e^{-\frac{1}{\tau_1}(t-s)} ds \\
g(n) &= \int_0^n \left[\dots \right] e^{-\frac{1}{\tau_1}(n-s)} ds \\
g(n) &= \sum_{i=0}^{n-1} \left[w(i) - \frac{1}{\tau_2} g(i-1) \right] e^{\frac{-1}{\tau_1}(n-i)}
\end{aligned}
$$

$$
\begin{aligned}
g(n) &= \int_0^n \left[w(s) - \frac{1}{\tau_2} g(s-1) \right] e^{-\frac{1}{\tau_1}(n-s)} ds \\
g(n) &= \sum_{i=0}^{n} \left[w(i) - \frac{1}{\tau_2} g(i-1) \right] e^{-\frac{1}{\tau_1}(n+1-i-1)} \\
g(n) &= e^{-\frac{1}{\tau_1}} \sum_{i=0}^{n} \left[w(i) - \frac{1}{\tau_2} g(i-1) \right] e^{-\frac{1}{\tau_1}(n-i-1)} \\
g(n) &= e^{-\frac{1}{\tau_1}} \left(\sum_{i=0}^{n} \left[w(i) - \frac{1}{\tau_2} g(i-1) \right] e^{-\frac{1}{\tau_1}(n-i-1)} + \left[w(n) - \frac{1}{\tau_2} g(n-1) \right] e^{\frac{1}{\tau_1}} \right) \\
g(n) &= e^{-\frac{1}{\tau_1}} \left[g(n-1) + w(n) - \frac{1}{\tau_2} g(n-1) e^{\frac{1}{\tau_1}} \right].
\end{aligned}
$$

C. Mathematical formulation of fitness-fatigue model combined to a kalman filter, according to Kolossa, Azhar, Rasche, et al. 2017

First, a linear system can be written as following:

$$\mathbf{x}_{n+1} = \mathbf{A}_n \mathbf{x}_n + \mathbf{B}_n \mathbf{w}_n + v_n \,,$$

with $\mathbf{x}_n$ a vector of p-dimensional states of the system. The system matrix $\mathbf{A}_n$ changes in state at each time, w_n denotes the system input and v_n a noise term that describes random changes in a given state. Since the states are not directly observable, we can estimate them by indirect measurements y_n such as

$$y_n = \mathbf{C}_n \mathbf{x}_n + o_n \,. \tag{.1}$$

Here, the output matrix $\mathbf{C}_n$ shows the influence of each state on the measurement y_n, and o_n denotes an observation noise.

If we consider a linear system with exponential decays, a time-invariant state matrix $\mathbf{A} = \mathbf{A}_n \forall n$, a zero initial state, the system state evolves with

$$x_n = \sum_{i=1}^{n-1} \mathbf{A}^{n-i-1} \mathbf{B}_i w_i \,.$$

Therefore, according to Kolossa, Azhar, Rasche, et al. 2017, this model is converted into the form a FFM in the sequel:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} , \mathbf{A}_n = A = \begin{pmatrix} e^{\frac{-1}{\tau_g}} & 0 \\ 0 & e^{\frac{-1}{\tau_h}} \end{pmatrix} , \mathbf{B}_n = \begin{pmatrix} e^{\frac{-1}{\tau_g}} \\ k_h(n) \cdot e^{\frac{-1}{\tau_h}} \end{pmatrix} , \mathbf{C}_n = \mathbf{C} = \begin{pmatrix} k_g & -1 \end{pmatrix} . \tag{.2}$$

Here, $\mathbf{x} = [x_1, x_2]^T$ is the state vector composed of fitness x_1 and fatigue x_2 states. $\mathbf{A}$ is the system matrix that contains exponential decay rates with parameters τ_1, τ_2 in the diagonal, $\mathbf{B}$ is the time-varying input matrix that also contains the two exponential decays and the weighting factor $k_h(n)$ (similar to $k_h(l)$ in Equation 1.29).

According to Equation .1 and neglecting the obervation noise o_n, we have

$$y_n = k_g \sum_{i=1}^{k-1} e^{\frac{-(n-i)}{\tau_g}} w_i - \sum_{i=1}^{k-1} k_h(i) e^{\frac{-(n-i}{\tau_h}} w_i \,. \tag{.3}$$

In addition to the linear system presented so far, KF uses noise covariance terms to describe random fluctuations in the state vector $\mathbf{x}_k$ and in the observation y_k. In the first step, the *a posteriori* state estimate $\hat{\mathbf{x}}_n$ is updated by the feedback

$$\hat{\mathbf{x}}_n = \mathbf{z}_n + \mathbf{K}_n(y_n - \mathbf{C}\mathbf{z}_n)\,,$$

where y_k is the measurement and $\mathbf{K}_n$ the optimal Kalman gain, such as

$$\mathbf{K}_n = \mathbf{M_n}\mathbf{C^T}(\mathbf{R} + \mathbf{C}\mathbf{M_n}\mathbf{C^T})^{-1}\,.$$

Here, $\mathbf{R}$ is the noise variance of the observation noise o_k. The matrix $\mathbf{M_n}$ is iteratively computed as

$$\mathbf{M_{n+1}} = \mathbf{Q} + \mathbf{A}\mathbf{M_n}\mathbf{A^T} - \mathbf{A}\mathbf{M_n}\mathbf{C^T}(R + \mathbf{C}\mathbf{M_n}\mathbf{C^T})^{-1}\mathbf{C}\mathbf{M_n}\mathbf{A^T}\,.$$

Therefore, a predicted *a priori* state estimate $\mathbf{z}_n$ is given by

$$\mathbf{z}_n = \mathbf{A}\hat{\mathbf{x}}_{n-1} + \mathbf{B}_{n-1}u_{n-1}\,.$$

D. Discretisation of model equations, taken from Busso 2017

Basically, the modelled performance on day i is defined by the balance between its production and removal on day $i-1$, such as

$$\hat{y}_i^{pharm} = \hat{y}_{i-1}^{pharm} e^{-k_{off}} + P_{i-1}\,,$$

where k_off denotes a constant for the removal of performance and P_{i-1} denotes the production of performance at day $i-1$. We consider $\hat{y}_0^{pharm}$ being equal to $\hat{y}_1^{pharm}$ and stationary. An offset term k_{on}^0 relates to the initial rate of removal with

$$\begin{aligned} k_{on}^0 &= \hat{y}_0^{pharm}(1 - e^{-k_{off}})\,, \quad \text{with} \\ P_0 &= k_{on}^0\,. \end{aligned}$$

In model T, a signal S for a day i is given by

$$S_i = S_{i-1} e^{-k_{out}^s - k_{on}^s} + w_i\,, \tag{.4}$$

where W_i denotes the measured training load for a given day, according to a discrete function $w(t)$. Accordingly, P_i becomes

$$P_i = k_{on}^0 + k_{on}^s \cdot S_i\,. \tag{.5}$$

In model TI, the inhibition I on day i takes the form

$$I_i = k_{in}^i \cdot w_i\,.$$

Therefore, from Equations .4 and .5 we have

$$S_i = S_{i-1} e^{-k_{out}^s - k_{on}^s \cdot (1 - I_i)} + w_i\,,$$

and

$$P_i = k_{on}^0 + k_{on}^s (1 - I_i) S_i\,.$$

In models TF and TIF, we have a fatigue term F such as

$$F_i = (k_{in}^f\, w_{i-1} + F_{i-1}) e^{k_{out}^f}\,,$$

with $F_0 = w_0 = 0$.

Therefore, $\hat{y}_i^{pharm}$ becomes a net performance $net\hat{y}_i^{pharm}$ with

$$net\hat{y}_i^{pharm} = \hat{y}_i^{pharm} - F_i\,.$$

E. Hierarchical mixed models used in Section 2.1

Section 2.1.4.1

$$
\begin{aligned}
\mathrm{Y}_i &\sim N\left(\mu, \sigma^2\right) \\
\mu &= \alpha_{j[i]} + \beta_1(\mathrm{condition}_2) + \beta_2(\mathrm{condition}_3) + \beta_3(\mathrm{velocity}) + \\
&\quad \beta_4(\mathrm{condition}_2 \times \mathrm{velocity}) + \beta_5(\mathrm{condition}_3 \times \mathrm{velocity}) \\
\alpha_j &\sim N\left(\mu_{\alpha_j}, \sigma^2_{\alpha_j}\right), \text{ for ID j} = 1, \ldots, \mathrm{J},
\end{aligned}
$$

where Y_i denotes either changes in $[lact_b]$ or $[cort_p]$ for each participant denoted ID.

Section 2.1.4.3

Rate of force development and impulsion

$$
\begin{aligned}
\mathrm{Y}_i &\sim N\left(\mu, \sigma^2\right) \\
\mu &= \alpha_{j[i]} + \beta_1(\mathrm{N}) + \beta_2(\mathrm{condition}_2) + \beta_3(\mathrm{condition}_3) + \\
&\quad \beta_4(\mathrm{condition}_2 \times \mathrm{N}) + \beta_5(\mathrm{condition}_3 \times \mathrm{N}) \\
\alpha_j &\sim N\left(\mu_{\alpha_j}, \sigma^2_{\alpha_j}\right), \text{ for ID j} = 1, \ldots, \mathrm{J},
\end{aligned}
$$

where, Y_i denotes either RFD_{peak}, RFD_{0-100} or impulsion for simplicity, N represents accumulated repetitions.

Surface electromyography analysis in time domain and torque

$$
\begin{aligned}
\mathrm{Y}_i &\sim N\left(\mu, \sigma^2\right) \\
\mu &= \alpha_{j[i]} + \beta_1(\mathrm{N}) + \beta_2(\mathrm{condition}_2) + \beta_3(\mathrm{condition}_3) + \\
&\quad \beta_4(\mathrm{velocity}) + \beta_5(\mathrm{condition}_2 \times \mathrm{velocity}) + \beta_6(\mathrm{condition}_3 \times \mathrm{velocity}) \\
\alpha_j &\sim N\left(\mu_{\alpha_j}, \sigma^2_{\alpha_j}\right), \text{ for ID j} = 1, \ldots, \mathrm{J},
\end{aligned}
$$

where Y_i denotes either the sum of EMG signals or normalised averaged torque, N represents accumulated repetitions.

Surface electromyography analysis in frequency domain: median frequencies

$$
\begin{aligned}
\mathrm{Y}_i &\sim N\left(\mu, \sigma^2\right) \\
\mu &= \alpha_{j[i]} + \beta_1(\mathrm{condition}_2) + \beta_2(\mathrm{condition}_3) + \beta_3(\mathrm{N}) + \\
&\quad \beta_4(\mathrm{condition}_2 \times \mathrm{N}) + \beta_5(\mathrm{condition}_3 \times \mathrm{N}) \\
\alpha_j &\sim N\left(\mu_{\alpha_j}, \sigma^2_{\alpha_j}\right), \text{ for ID j} = 1, \ldots, \mathrm{J},
\end{aligned}
$$

where Y_i denotes either the sum of MDF computed from power spectrum, N represents accumulated repetitions.

Surface electromyography analysis in frequency domain: rate decay of median frequencies

$$
\begin{aligned}
\mathrm{y}_i &\sim N\left(\mu, \sigma^2\right) \\
\mu &= \alpha_{j[i]} + \beta_1(\mathrm{condition}_2) + \beta_2(\mathrm{condition}_3) + \quad \beta_3(\mathrm{condition}_2 \times \mathrm{S}) + \\
&\beta_4(\mathrm{condition}_3 \times \mathrm{S}) \\
\alpha_j &\sim N\left(\mu_{\alpha_j}, \sigma^2_{\alpha_j}\right), \text{ for ID j} = 1, \ldots, \mathrm{J},
\end{aligned}
$$

where y_i denotes either the rate decay of median frequencies, S represents the set number.

F. Appendix: Training load quantification in Short track speed-skating

Specific training

On-ice sessions refer to specific training. Session TLs were calculated from Power Output (PO, W), volume and ice properties. Individual PO depends on power required to change kinetic energy (P_{kin}), power required to overcome air and ice resistance (P_{aero} and P_{ice}, respectively). Let us define

$$\begin{aligned} P_{kin} &= \frac{\frac{1}{2}(m v_f^2 - v_i^2)}{t}, \\ P_{aer} &= \frac{1}{2} A C_D \, \rho \, v^3 \quad \text{and} \\ P_{ice} &= C_f \, m \, g \, \bar{v}. \end{aligned}$$

In this context, m denotes the mass of the athlete and that of the equipment, v_f is the maximal velocity reached during the run, v_i is the initial velocity being null, $\bar{v}$ is the mean velocity and t is the exercise duration. The effective frontal area AC_D is a standardised fixed value of $0.25\,\mathrm{m}^2$ according to subjects corpulence and Van Ingen Schenau[56]. Also, ρ denotes the air density recorded at 1850 meters above sea and is equal to $1.029\,\mathrm{kg.m}^{-3}$. The friction coefficient C_f is standardised as $C_f = 0.006$, according to maximal values found by De Koning *et al.*[57] and due to a track with sharper turns. Finally, g denotes the acceleration due to the gravity, equal to $9.80665\,\mathrm{m.s}^{-2}$.

Thus,

$$PO = P_{kin} + P_{aer} + P_{ice}.$$

Relative intensity of the session (I_{ice}, as a percentage of the maximal PO) can now be determined as

$$I_{ice} = \frac{I_{ice}^f \, N_f + I_{ice}^b \, (N - N_f)}{N}. \tag{S1}$$

This relative exercise intensity includes both forward and backward positions denoted I_{ice}^f and I_{ice}^b respectively, with

$$\begin{aligned} I_{ice}^f &= \frac{PO}{\max PO} + C \, , \\ I_{ice}^b &= I_{ice}^f - E \, I_{ice}^f \, . \end{aligned}$$

Here, C denotes the ice impact on skating for an ice quality (Q_{ice}) arbitrary

measured by athletes on a 0-10 Borg scale and averaged. If Q_{ice} is below 7.5 arbitrary units (a.u), a linear penalisation is attributed such as $C = -0.008\, Q_{ice} + 0.06$, where α and β coefficients were estimated from at least two equal performances with different values of Q_{ice}. In addition, E denotes the skating economy due to drafting and N denotes the overall number of laps with also a distinction for the forward position (N_f). Finally, ice session training load is

$$TL_{ice} = I_{ice}\, V\, K\, (\frac{I_{RPE}}{\max I_{RPE}})\, \rho, \tag{S2}$$

where V is the volume parameter defined as the product of the number of laps run and the distance of a lap; K depends on the subject's gender with $K = 0.64\, e^{1.92\, I}$ for males and $K = 0.86\, e^{1.67\, I}$ for females respectively and according to Banister, Good, Holman, et al. 1986; I_{RPE} is the rate of perceived exertion quoted on a 6-20 Borg scale, $\max I_{RPE}$ is the maximal value that can be quoted ($\max I_{RPE} = 20$); ρ denotes the density parameter, such as $\rho = \frac{1}{2}\, \rho_s$ with ρ_s the density of the session (%) which represents the effective work done by the athlete.

Non-specific training

Training loads of resistance training (TL_{RT}), aerobic training (TL_{aer}), repeated sprint training (TL_{RS}) and activation sessions (TL_{act}, specific warm-up) were also quantified as

$$TL_{RT} = I_{RT}\, V\, K\, (\frac{I_{RPE}}{\max I_{RPE}})\, \rho, \tag{S3}$$

$$TL_{aer} = I_{RPE}\, T\, K\, \rho_s\, k_{aer}, \tag{S4}$$

$$TL_{RS} = I_{RS}\, V\, K\, (\frac{I_{RPE}}{\max I_{RPE}})\, \rho \quad \text{and} \tag{S5}$$

$$TL_{act} = I_{RPE}\, T\, K\, \rho_s\, k_{off}. \tag{S6}$$

Here I_{RT} denotes the intensity in percentage of the maximal repetition, V is the volume defined by the number of repetitions, T is the total time of exertion, k_{aer} and k_{off} denote a weighting factor for aerobic and activation exercises such as $k_{aer} = 5$ a.u (empirically defined by the coach) and $k_{off} = 15$ a.u respectively. Any of the training sessions are weighted by I_{RPE}. However, a specific intensity was only quantifiable for TL_{RT} and TL_{RS} and further considered in the training load calculation.

According to the training condition, Equations S2 – S6 respectively define the discrete function $w(t)$.

G. Appendix: Fitted parameters using elastic net regularisation and principal component regression

In the sequel, we provide the fitted parameters of model concerned in Figure 3.7.

G.1. Elastic net regularisation

Parameter	Coefficient
Intercept	16.84
past_perf	0.47
S4	0.11
S6	0.06
ser_density	0.02
imp_load	0.01
imp_density	0.01
imp_relativeInt	0.00
imp_avgPower	0.00
imp_RPE	0.00
imp_session_duration	0.00
ser_load	0.00
ser_avgPower	0.00
ser_maxPower	0.00
ser_RPE	0.00
ser_relativeInt	0.00
ser_session_duration	0.00
ice_quality	-0.00
rest_days	-0.01
imp_maxPower	-0.03
S7	-0.15
S3	-0.18
S2	-0.22
S5	-0.23

G.2. Principal component regression: contribution of variables per dimension

Independent variable	contribution (%)	Dimension
ser_load	10.07	1
imp_relativeInt	9.99	1
ser_RPE	9.75	1
ser_relativeInt	9.70	1
ser_maxPower	9.33	1
imp_density	28.21	2
imp_session_duration	22.20	2
imp_RPE	14.24	2
ser_density	10.01	2
ser_session_duration	4.61	2
rest_days	28.70	3
ser_density	14.68	3
ice_quality	11.97	3
imp_load	11.19	3
imp_RPE	10.20	3

H. Appendix: Fitted parameters using elastic net regularisation and principal component regression

Statistical modelling in a small data set context can lead to statistical power issues and may suffer from biased parameters estimation. To tackle this issue, the modelling was conduced in a Bayesian framework. We counterbalanced the lack of data (participants) by providing a priori information inside the models, based on empirical knowledge and literature. The Hamilton Monte Carlo algorithm was used to infer the parameters of models and caution has been taken to diagnose their convergence (Vehtari, Gelman, Simpson, et al. 2019). To figure out the relevance of variable inclusion in the models and to provide an alternative to significant testing of the null hypothesis (H_0), we computed Bayes Factors (BF_{10}). Such a factor represents a continuous measure of evidence for the alternative hypothesis (H_1) over H_0. Based on theory of Jeffreys 1961 and according to guidelines of Lee and Wagenmakers 2014, we provided the following classification for interpretation : $BF_{10} \geqslant 100$, 30–100, 10–30, 3–10, 1–3 correspond to an extreme, very strong, strong, moderate and anecdotal evidence for H_1, respectively. A BF_{10} of 1 means there is no evidence of an hypothesis over the other. Below this value, the evidence is against H_1 or for H_0 following the inverse of the mentioned scale.

H.1. Reference Measures, section 4.2.2

To ensure that reference measures were valid, LMM was computed to evaluate the relationship between the Cm and speed calculated from force platform measurements. Speed and participants were settled as fixed and random effects respectively, in order to consider the variability of Cm among participants. Relationship between the portable metabolic system-derived variables and the force platform-derived variables were also assessed by computing a LMM where the mechanical power and participants corresponded to fixed and random effects, respectively. For each LMM, an ICC was reported to highlight the fraction of the total variance in the data accounted for between-subject variation. It justifies the inclusion of participants as random effects in the model. Finally, linear models were used to examine the mean relationships between (i) GCT and running mechanics (running speed and frequency), (ii) leg stiffness and running mechanics.

H.1.1. Mechanical cost of running and speed relationship

$$y_{ij} = \beta_0 + S_{0i} + \beta_1 speed_j + \epsilon_{ij}$$

with y_{ij} being a Cm value for a subject i, $speed_j$. β_0 denotes the fixed effect intercept, S_{0i} is the offset term intercept which represents the deviation from β_0 for the subject i. β_1 is the parameter for the speed predictor in which a weakly informative prior $\beta_1 \sim \mathcal{N}(0, 10)$ was assigned and ϵ_{ijk} is the observation-level error.

H.1.2. Relationships between consumed metabolic energy and external mechanical power

$$y_{ij} = \beta_0 + S_{0i} + (\beta_1 + S_{1i})\, \dot{W}ext_j + \epsilon_{ij}\ , \tag{.6}$$

with y_{ij} being the $\dot{\mathrm{V}}\mathrm{O}_2$ for a subject i and $\dot{W}ext_j$ (*i.e.* the external mechanical power). β_0 denotes the fixed effect intercept, S_{0i} is the offset term intercept which represents the deviation from β_0 for the subject i, β_1 is the parameter for the $\dot{W}ext$ predictor with a weakly informative prior $\beta_1 \sim \mathcal{N}(0, 10)$ assigned , S_{1i} is the random slope for each subject and ϵ_{ij} is the observation-level error.

H.1.3. Relationships between GCT, LSS and stride mechanics (running speed and frequency)

$$y_i = \beta_0 + \beta_n x_i + \epsilon_i,$$

with yi the variable response GCT or LSS for a subject i, β_0 the intercept, β_n the parameter for the predictor x (running speed or frequency) and ϵi the error term. A weakly informative prior $\beta_n \sim \mathcal{N}(0, 10)$ was assigned to both running speed and running frequency predictors.

H.2. Comparisons between Stryd and reference measures

H.2.1. Mechanical Power, sections 4.2.3.1 and 4.2.3.2

To assess the linear relationship between the Stryd mechanical power and the metabolic energy consumption during the MAS test, we defined the following model,

$$y_i = \beta_0 + \beta_1 \dot{W}ext_i + \epsilon_i, \tag{.7}$$

with y_i being the $\dot{\mathrm{V}}\mathrm{O}_2$, β_0 denoting the intercept, $\beta_1 \sim \mathcal{N}(0, 10)$ being a weakly informative prior and ϵ_i the error term.

In addition to coefficient of determination and 95% CI, correlation between variables were observed through pair-wise Bayesian correlation tests using non-informative Jeffrey's priors (Jeffreys 1961). A similar linear model (Equation (.7)) was computed to assess the relationship between PO measured by the force platform and PO assessed from the power meter.

Thereafter, the effects of the system of measurement (force platform and Stryd power meter), the speed and the participant were evaluated. Due to random measurement errors and technological issues (e.g., mismatch between foot strikes and the force platform area), it has been required to deal with missing data as well as different number of repeats between systems. Consequently, a LMM was preferred to a repeated measures analysis of variance. Through a design-driven approach (Barr, Levy, Scheepers, et al. 2013), subject varying intercept and slopes were included in addition to fixed effects. This model allowed us to consider the inter-subject and intra-subject variability (e.g., heterogeneous PO levels and individual PO kinetics in response to speed and strides changes). Independent variables were standardised prior modelling, easing the interpretation by allowing the direct comparisons of estimated parameters. The model was defined as,

$$\begin{aligned} y_{ijk} = \beta_0 + S_{0i} + \beta_1\, device_k \\ + (\beta_2 + S_{1i})\, speed_j \\ + \beta_3(speed_j\, device_k) + \epsilon_{ijk}\ , \end{aligned} \tag{.8}$$

with y_{ijk} being the response variable for a subject i, speed j and device respectively k. In this equation, β_0 denotes the fixed effect intercept, S_{0i} is the offset term intercept which represents the deviation from β_0 for the subject i, β_n are the parameters for each corresponding predictor, S_{1i} is the random slope for each subject and ϵ_{ijk} is the observation-level error. Priors were chosen according to empirical knowledge and literature. Because the relationship between devices (Stryd power meter and force plateforms) remains unknown but presuming an underestimate of the Stryd power meter, a vague prior was fixed such as $\beta_0 \sim \mathcal{T}(3, 0, 10)$ and $\beta_1 \sim \mathcal{N}(0, 1000)$. According to the well-known strong and positive relationship between running speed and external mechanical power (Chelly and Denis 2001), a weakly but more informative prior (i.e. with a lower variance) was fixed to the speed parameter with $\beta_2 \sim \mathcal{N}(0, 200)$.

H.2.2. Ground Contact Time, section 4.2.3.3

In the same way, effects of the MAS test parameters on GCT were assessed through the Bayesian LMM (Equation (.8)). Weakly informative priors were assigned to the parameters of the force platforms, power meter and speed. This choice was motivated from empirical knowledge and previous findings about changes in contact time with velocity (Blickhan 1989): $\beta_0 \sim \mathcal{T}(3, 0, 10)$, $\beta_1 \sim \mathcal{N}(0, 1)$, $\beta_2 \sim \mathcal{N}(0, 1)$.

H.2.3. Leg Stiffness, section 4.2.3.4

The last variable of interest was analysed following the same procedure (Equation (.8)). Weakly informative priors were chosen according to the trust in leg stiffness

values (Farley and Gonzalez 1996) in order to estimate the posterior distributions for each parameters as $\beta_0 \sim \mathcal{T}(3, 0, 10)$, $\beta_1 \sim \mathcal{N}(0, 10)$, $\beta_2 \sim \mathcal{N}(0, 1)$.