Note. This article will be published in a forthcoming issue of the *International Journal of Sports Physiology and Performance*. The article appears here in its accepted, peer-reviewed form, as it was provided by the submitting author. It has not been copyedited, proofread, or formatted by the publisher.

Section: Original Investigation

Article Title: Relationships Between the External and Internal Training Load in Professional Soccer: What Can We Learn From Machine Learning?

Authors: Arne Jaspers[1], Tim Op De Beéck[2], Michel S. Brink[3], Wouter G.P. Frencken[3], Filip Staes[4], Jesse J. Davis[2], and Werner F. Helsen[1]

Affiliations: [1]Department of Movement Sciences, Movement Control and Neuroplasticity Research Group; [2]Department of Computer Science; KU Leuven, Leuven, Belgium. [3]Center for Human Movement Sciences, University of Groningen, University Medical Center, Groningen, The Netherlands. [4]Department of Rehabilitation Sciences, Musculoskeletal Rehabilitation Research Group, KU Leuven, Leuven, Belgium.

Journal: *International Journal of Sports Physiology and Performance*

Acceptance Date: December 12, 2017

DOI: https://doi.org/10.1123/ijspp.2017-0299

"Relationships Between the External and Internal Training Load in Professional Soccer: What Can We Learn From Machine Learning?" by Jaspers A et al.
International Journal of Sports Physiology and Performance

Abstract

Purpose: Machine learning may contribute to understanding the relationship between the external load and internal load in professional soccer. Therefore, the relationship between external load indicators and the rating of perceived exertion (RPE) was examined using machine learning techniques on a group and individual level. **Methods:** Training data were collected from 38 professional soccer players over two seasons. The external load was measured using global positioning system technology and accelerometry. The internal load was obtained using the RPE. Predictive models were constructed using two machine learning techniques, artificial neural networks (ANNs) and least absolute shrinkage and selection operator (LASSO), and one naive baseline method. The predictions were based on a large set of external load indicators. Using each technique, one group model involving all players and one individual model for each player was constructed. These models' performance on predicting the reported RPE values for future training sessions was compared to the naive baseline's performance. **Results:** Both the ANN and LASSO models outperformed the baseline. Additionally, the LASSO model made more accurate predictions for the RPE than the ANN model. Furthermore, decelerations were identified as important external load indicators. Regardless of the applied machine learning technique, the group models resulted in equivalent or better predictions for the reported RPE values than the individual models. **Conclusions:** Machine learning techniques may have added value in predicting the RPE for future sessions to optimize training design and evaluation. Additionally, these techniques may be used in conjunction with expert knowledge to select key external load indicators for load monitoring.

Keywords: Football, athlete monitoring, global positioning system, rating of perceived exertion, predictive modelling

Introduction

Nowadays, professional soccer clubs monitor training and match load to optimize physical fitness and reduce injury risk.[1] When considering training and match loads, it is typical to distinguish between the external and internal load.[2] The external load represents the dose performed and the internal load represents the psychophysiological stress experienced by the player.[2] The external load is generally defined as all locomotor and non-locomotor activities performed by players.[2 3] Global positioning systems (GPS) and inertial sensors are used for monitoring external load indicators (ELIs) such as the distance covered and jumps.[3] The internal load can be quantified using the rating of perceived exertion (RPE), which is often considered a good indicator of the global internal load.[4] Due to differences in individual characteristics (e.g., training history and actual physical fitness), similar external loads can result in different internal loads for players. Insights into the relationship between the external and internal load can improve load management and help to optimize physical fitness and support injury prevention.[5]

To date, several studies about team sports have focused on the relationship between the external and internal load. In these studies, the data were analyzed using traditional statistical methods such as Pearson correlation coefficients, multiple regression and general linear models with partial correlation coefficients.[6-8] Recently, a study in Australian football (AFL) found that artificial neural networks (ANNs), a machine learning approach, more accurately predicted the RPE in response to ELIs compared to traditional statistics.[9] Other machine learning techniques could be used for this task as well, and each technique has strengths and weaknesses.[10]

In general, the data-driven approach of machine learning is able to capture linear and non-linear relationships between various ELIs and the response variable RPE.[10] Given a large set of ELIs, machine learning approaches can automatically identify the specific ELIs that are most predictive of the RPE, often without correcting for multicollinearity or using expert knowledge to hand select a set of ELIs. This can aid in evaluating newly developed external load metrics that come with improved tracking systems such as GPS technology and inertial movement sensors.[11]

Another advantage of machine learning is its ability to detect possible inter-player differences. In the AFL study using machine learning techniques, various ELIs were examined to determine their predictive value for each player’s RPE.[9] Inter-player differences were found for ELIs and their contribution to an individual’s RPE.[9] For most players, the distance covered was the most predictive ELI for the RPE. However, for some players, the distance covered per minute or distance covered at high-speed (>14.4 km.h^{-1}) had a higher predictive value, indicating that individual differences should be considered when evaluating dose and response to training load.[9]

Even though AFL and soccer are both running-based team sports, each sport imposes different physical demands on players due to differences in rules, pitch dimensions, player rotations versus substitutions, and playing time.[12] In comparison to soccer players, AFL players typically cover 2.6 times greater distance (1322m versus 517m) at very high-speed (19.8-25.1 km.h^{-1}) and 3.5 times greater distance (328m versus 93m) at sprinting speed (>25.2 km.h^{-1}) in matches.[12] When comparing the absolute number of maximal acceleration efforts (>2.78 m.s-2) to the absolute number of high-speed efforts (19.8-25.1 km.h-1), AFL players show a 1:1 ratio whereas soccer players exhibit a ratio of 1.7, indicating that numerous accelerations during matches do not result in high-speed efforts.[12] Based on this comparison, it may be unlikely that the results regarding the most predictive ELIs and inter-player differences in AFL will generalize to professional soccer. To our knowledge, no prior study in professional soccer has investigated the relationship between ELIs and RPE using machine learning techniques to determine which ELIs are most predictive of the RPE or to examine possible inter-player differences.

In summary, the current study aims to evaluate the ability of machine learning techniques to (1) predict the RPE from a given set of ELIs; (2) identify which ELIs for soccer players contribute most to the RPE; and (3) evaluate both group and individual models to examine possible inter-player differences regarding the relationship between ELIs and RPE.

Methods

Subjects

Data from 38 professional soccer players (22.7 ± 3.4 years, 1.83 ± 0.06 m, 77.0 ± 6.7 kg, and 10.3 ± 1.8% body fat) competing for a team in the highest league in the Netherlands were included. Goalkeepers' data were excluded from the study due to different physical demands. The study was conducted according to the requirements of the Declaration of Helsinki and was approved by the KU Leuven ethics committee (file number: s57732).

Design

Data were collected from pre-season and in-season training sessions over two seasons (2014-2015 and 2015-2016). Like with Bartlett et al.,[9] this study focused on the relationship between ELIs and RPE in training sessions. Therefore, data from matches, on-field recovery sessions, and rehabilitation sessions were excluded from the analysis. For each training session, the external load was measured using 10 Hz GPS and 100 Hz accelerometer technology (Optimeye S5, Catapult Sports, Melbourne) in accordance with the recommendations for collecting and processing GPS data in sports.[11] The internal load was measured using the RPE. Each player reported his RPE approximately 30 minutes after the training session using the modified Borg CR-10 scale.[13] All players were familiarized with the use of RPE before the beginning of the study and were instructed to rate their perceived effort for the whole training session.[4] Furthermore, each player was asked in isolation for his RPE to minimize the influence of factors such as peer pressure.[14]

The first season contained data from 23 players. The number of sessions recorded per player ranges from 35 to 160 with a mean and standard deviation of 125 ± 34 sessions. The second season contained data from 28 players. The number of sessions recorded per player ranged from 51 to 163 with a mean and standard deviation of 109 ± 33 sessions. As players frequently switch teams in professional football, only 13 players appeared in both seasons.

Methodology

To examine the relationship between the external load and RPE using machine learning, a set of 67 ELIs that could be exported from the manufacturer's software (Sprint™ version 5.1.7, Catapult Sports, Melbourne, Australia) was selected to capture the external load of a training session. The set of ELIs can be divided into high-level categories about duration, distance, speed, acceleration and deceleration, PlayerLoad (i.e., a metric based on accelerometry),[15] and repeated high-intensity effort (RHIE) activity (Table 1). The first goal was to identify the ELIs that are most predictive of the RPE. Therefore, a model was constructed that accurately predicts what a player's reported RPE (internal load) will be based on the observed value for all ELIs in a training session.

The mean absolute error (MAE) was used to assess a model's predictive performance. This metric calculates the mean of the absolute errors (i.e., |(reported RPE value) – (predicted RPE value)|) over all predictions. The MAE is easy to interpret as it uses the same unit as the RPE value: a MAE of 1 means that, on average, the predicted RPE is one value below or above the reported RPE. While a MAE of zero is unrealistic, the goal is to minimize a model's MAE.

To construct predictive models, two standard machine learning techniques were considered as well as one naive baseline method:

Artificial neural networks (ANN)

ANNs are a standard approach for constructing non-linear models that often exhibit good predictive performance.[10] However, a disadvantage of ANNs is that the resulting models are difficult to interpret (i.e., they do not provide insight into the interactions that are modelled among ELIs).

Least absolute shrinkage and selection operator (LASSO)

This technique is an advanced version of linear regression.[16] When setting the regression coefficients, LASSO contains a mechanism that biases many of them to be zero. Consequently, LASSO only selects a subset of the ELIs, those with a non-zero coefficient, to be included in the model. This results in both better interpretability and more robustness to multicollinearity among the input variables than traditional linear regression. As LASSO constructs a linear model, it is more robust to small sample sizes compared to the more expressive ANNs.

Additionally, a well-known LASSO-based approach can be used to compute importance scores of the ELIs.[17] The importance scores are calculated as the probability that an ELI is selected by the LASSO model and fall in the range of zero to one. Higher scores denote more important ELIs. In general, the presence of collinearity among the input ELIs tends to result in lower importance scores.

Baseline

This model does not consider the external load and always predicts the average RPE value over all training sessions used to construct the model. This model assumes that none of the ELIs are predictive of the RPE. While a MAE of zero is a lower bound (i.e., a perfect predictive model), the baseline provides a realistic upper bound for the MAE. A valuable predictive model should have a lower MAE than this baseline.

Data analysis

Two experiments were conducted. Each one employed standard machine learning methodology and subdivided the data into two disjoint sets: the learning set and testing set. Each machine learning approach used the data in the learning set to construct a model. The independent testing set was used to estimate a model's predictive performance on unseen (that is, future) data. Specifically, each model made a prediction for the reported RPE associated with every training sessions in the testing set, and the MAE was computed for these predictions. In addition, 90% confidence intervals (CI) and effect sizes were calculated.[18 19]

The first experiment evaluated the value of group models. The temporal nature of the data was preserved by partitioning the data based on seasons: data from the first season served as the learning set and the data from the second season as the testing set. A consequence of the seasonal split was that each model made predictions for unseen players, that is, players who had no data in the learning set. One group model was constructed using each learning approach. The most predictive ELIs were identified by inspecting the most accurate learned model.

The second experiment examined the impact of accounting for inter-player differences. As only a few players appeared in both seasons, there was insufficient data to consider season-based partitioning of the data. Therefore, season 1 and season 2 were treated separately. Each season's data was subdivided temporally such that the first 75% served as the learning set and the last 25% served as the testing set. Using each learning approach, both one group model and an individual model for each player was constructed. The group model was constructed using data from all the players in the learning set. An individual model for each player was constructed by only considering that specific player's training session data in the learning set. A global mean of the absolute errors of all individual models was calculated so that the metric aligned with how the group model's MAE was computed.

For automated preprocessing and advanced analysis, custom Python scripts were developed using Python Pandas for data handling[20] and Sklearn for machine learning.[21]

Results

The average RPE for all 5917 analyzed training sessions was 3.59 ± 1.46 AU. The following descriptive statistics were calculated for these commonly reported ELIs: duration 70 ± 16 minutes, total distance covered 4614 ± 1576 m, distance covered at high-speed (>15 km.h^{-1}) 426 ± 351 m, and total distance covered per minute 65 ± 14 m.min^{-1}.

Table 2 shows the MAEs and 90% CIs for the group models constructed using the data from season 1 and evaluated on the data from season 2. In addition, the effect sizes are shown for the MAEs of ANN and LASSO group models compared to the baseline's MAE. Both the ANN and LASSO models outperform the baseline. Compared to the baseline, the LASSO model resulted in a 29.8% reduction in the MAE when predicting the RPE of unseen training sessions from season 2. Moreover, the LASSO model made more accurate predictions than the ANN model. A trivial effect size was found for ANNs compared to the baseline, while a small effect size was found for the LASSO group model compared to the baseline.

Table 3 displays the ELIs, and their corresponding importance scores, selected by the LASSO group model (learned on the data from season 1) that most contribute to predicting the RPE.

Table 4 reports the MAEs and 90% CIs for individual and group models that were constructed and evaluated on season 1 and season 2 separately. Additionally, the effects sizes are presented for the comparison of the MAEs of ANN and LASSO models (i.e., both individual and group models) with the baseline. In all eight cases, the learned models had a lower MAE score than the baseline. Regardless of learning method, the group models resulted in equivalent or even more accurate predictions of the reported RPE values than the individual models.

Discussion

This study aimed to evaluate the ability of machine learning techniques to predict the RPE of soccer training sessions from a set of ELIs. Additionally, it aimed to identify the ELIs which are most predictive of RPE within a professional soccer context. Finally, it attempted to explore inter-player differences for how ELIs contribute to each player's RPE.

The constructed ANN and LASSO models outperformed the baseline indicating that it is possible to construct machine learning models that capture a part of the relationship between ELIs and RPE in professional soccer. Additionally, it suggests that a good strategy is to start with a large set of ELIs, as opposed to hand selecting a small number of ELIs to reduce the chance of discarding a relevant ELI. Moreover, a strength of machine learning techniques is their ability to automatically select a subset of predictive ELIs, often without correcting for multicollinearity. Therefore, this method may provide new insights and support expert knowledge in the selection of key load indicators for monitoring strategies.

The LASSO technique identified various ELIs as contributing the most to the perceived exertion in professional soccer (Table 3). These ELIs are partly in line with earlier findings in professional soccer using a smaller set of ELIs.[6 8] However, as GPS devices from different manufacturers are used in the other studies, it is difficult to compare findings.[11]

The novel important ELIs are indicators regarding decelerations. The results of this study indicate that this type of load, next to other ELIs, may contribute to a player's RPE. Previously, mainly concentric, energy-demanding efforts were associated with higher RPE values in professional soccer.[6 8] Decelerating efforts are related to eccentric activity.[22] This type of muscle activity has a lower energy cost in comparison with concentric muscle activity.[23] However, this type of eccentric contractions might more easily induce muscle damage.[22 23] Therefore, monitoring ELIs concerning decelerations can be particularly important.

Both individual and group models captured part of the relationship between ELIs and RPE. In contrast to Bartlett et al.,[9] we found that group models using ANN and LASSO techniques demonstrate an equivalent or superior accuracy for both season 1 and 2 compared to individual models when predicting RPE based on ELIs. A combination of diverse underlying factors may explain these results.

First, these findings are in contrast to the theoretical model of Impellizzeri et al.,[2] which states that the internal load (RPE) results from the interaction between the external load (ELIs) and individual characteristics. The results of our study may indicate that there is less variation in the external loads and individual characteristics of professional soccer players than in AFL so there is less impact on the reported RPE. It is possible that there are greater differences in positional activity profiles[12 24] and in individual characteristics (e.g., body composition and aerobic capacity) in AFL compared to professional soccer, which result in a more heterogenous group in AFL. The descriptive statistics for the ELIs and RPE clearly exhibit lower average values and less variation for professional soccer training sessions compared to AFL training sessions.[9] These inter-sport differences may partly explain the results indicating the presence of other ELIs that mutually determine the RPE for (most of) the players within a professional soccer team.

On the other hand, the sample size (i.e., the number of data points used to construct the model) is another factor which may have contributed to the equivalent performance of the group models. The group models are learned using a much larger sample size of more than 2000 data points compared to the individual models which typically relied on less than 100 data points. Nonetheless, we find that individual models constructed with the LASSO method perform similarly to the group models as the technique is robust to small sample sizes. If more data were available for each player, we would expect the individual models’ performance to improve. However, from a practical perspective this does not seem realistic. In professional soccer, only 100-150 training sessions (i.e., data points) are conducted per season per player. Additionally, players are often transferred which makes it difficult to obtain data over multiple seasons.

The current study focused on the relationship between ELIs and RPE for training sessions and matches were thus excluded. In future research, the same method could be applied to examine if similar ELIs influence the RPE for matches, or if different ELIs determine the RPE values of matches. However, as mentioned, machine learning requires sufficient amounts of data to build accurate predictive models. This could be a limitation due to the relative small number of games in a season. Additionally, the RPE for matches may be influenced by contextual factors.[25]

Recently, the differential RPE (dRPE) has demonstrated its added value by quantifying respiratory and muscular perceived exertion.[26-28] Using the dRPE may further clarify if specific ELIs have a higher impact on central (i.e., breathlessness) or local (i.e., leg muscle exertion) perceived exertion. These insights can aid in optimizing load and adaptation in terms of physiological (i.e., cardiorespiratory system) and biomechanical (i.e., musculoskeletal system) pathways.[29] Additionally, measures of recovery and psychosocial factors were not considered. Therefore, the inclusion of measures such as pre-training perceived wellness and recovery may further clarify the RPE outcome for a given external training load.[30 31]

The identification of key ELIs may aid in the evaluation of players' training dose and response over time using efficiency ratios (i.e., the proportion between RPE and ELIs).[32 33] For example, some ELIs may be perceived as less exerting at the end of preseason or a rehabilitation process compared to the beginning due to improvements in physical fitness. Consequently, a consistent deviation between the expected and reported RPE may be used as an efficiency ratio. This ratio could be used to exhibit if players evolve over time in their ability to deal with the external load. However, further research is needed regarding efficiency ratios relating to changes in fitness or fatigue.

Practical applications

Machine learning techniques may have added value in predicting the RPE for future training sessions and in selecting key ELIs for load monitoring in professional soccer. This study identified novel ELIs that should be considered such as high-magnitude decelerations that contribute to the RPE.

In addition, group models may have an added value in predicting the RPE for individual players: they can be applied to any player whereas an individual model is only applicable to that specific player. Hence, group models can make predictions for newly transferred or youth players, for whom there is often little (or no) available data. From a monitoring perspective, a dashboard for player monitoring may initially be made with similar ELIs for the players within a team. In case more data is available for a specific player, an individual model can be constructed and a customized dashboard can be monitored.

Conclusion

Our study confirmed that machine learning techniques are able to predict RPE based on a large set of ELIs collected during two seasons in professional soccer. Secondly, these techniques can be applied to support expert knowledge for the selection of key ELIs such as decelerations and, accordingly, improve load management strategies. Lastly, group models predicted the RPE with an equivalent or even better accuracy than individual models. Possible limitations of the applied machine learning approaches were discussed. In addition, guidelines for future machine learning research and practical applications were provided.

Acknowledgements

The authors would like to thank the players and both physical and medical staff for their participation in this study.

Arne Jaspers is supported by a research grant from the Agency for Innovation by Science and Technology–IWT, Belgium (IWT 130841). Jesse Davis and Tim Op De Beeck are partially supported by the KU Leuven Research Fund (C22/15/015) and the Interreg V A project NANO4Sports. The authors have no conflicts of interest to declare.

References

1. Akenhead R, Nassis GP. Training load and player monitoring in high-level football: current practice and perceptions. *Int J Sports Physiol Perform* 2016;11(5):587-93. doi: 10.1123/ijspp.2015-0331

2. Impellizzeri FM, Rampinini E, Marcora SM. Physiological assessment of aerobic training in soccer. *J Sports Sci* 2005;23(6):583-92. doi: 10.1080/02640410400021278

3. Gabbett TJ. The training-injury prevention paradox: should athletes be training smarter and harder? *Br J Sports Med* 2016;50(5):273-80. doi: 10.1136/bjsports-2015-095788

4. Impellizzeri FM, Rampinini E, Coutts AJ, et al. Use of RPE-based training load in soccer. *Med Sci Sports Exerc* 2004;36(6):1042-7. doi: 10.1249/01.mss.0000128199.23901.2f

5. Drew MK, Cook J, Finch CF. Sports-related workload and injury risk: simply knowing the risks will not prevent injuries. *Br J Sports Med* 2016:1-4. doi: 10.1136/bjsports-2015-095871 [published Online First: 10 May 2016]

6. Gaudino P, Iaia FM, Strudwick AJ, et al. Factors influencing perception of effort (session rating of perceived exertion) during elite soccer training. *Int J Sports Physiol Perform* 2015;10(7):860-4. doi: 10.1123/ijspp.2014-0518

7. Lovell TW, Sirotic AC, Impellizzeri FM, et al. Factors affecting perception of effort (session rating of perceived exertion) during rugby league training. *Int J Sports Physiol Perform* 2013;8(1):62-9.

8. Scott BR, Lockie RG, Knight TJ, et al. A comparison of methods to quantify the in-season training load of professional soccer players. *Int J Sports Physiol Perform* 2013;8(2):195-202. doi: 10.1123/ijspp.8.2.195

9. Bartlett JD, O'Connor F, Pitchford N, et al. Relationships between internal and external training load in team sports athletes: evidence for an individualised approach. *Int J Sports Physiol Perform* 2017;12(2):230-4. doi: 10.1123/ijspp.2015-0791

10. Bishop C. Pattern recognition and machine learning (Information science and statistics), 1st edn. 2006. corr. 2nd printing edn: Springer, New York, 2007.

11. Malone JJ, Lovell R, Varley MC, et al. Unpacking the black box: applications and considerations for using GPS devices in sport. *Int J Sports Physiol Perform* 2017;12(Suppl 2):S218-26. doi: 10.1123/ijspp.2016-023

12. Varley MC, Gabbett T, Aughey RJ. Activity profiles of professional soccer, rugby league and Australian football match play. *J Sports Sci* 2014;32(20):1858-66. doi: 10.1080/02640414.2013.823227

13. Foster C, Florhaug JA, Franklin J, et al. A new approach to monitoring exercise training. *J Strength Cond Res* 2001;15(1):109-15. doi: 10.1519/00124278-200102000-00019

14. Malone JJ, Di Michele R, Morgans R, et al. Seasonal training-load quantification in elite English premier league soccer players. *Int J Sports Physiol Perform* 2015;10(4):489-97. doi: 10.1123/ijspp.2014-0352

15. Barrett S, Midgley A, Lovell R. PlayerLoad™: reliability, convergent validity, and influence of unit position during treadmill running. *Int J Sports Physiol Perform* 2014;9(6):945-52. doi: 10.1123/ijspp.2013-0418

16. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Series B Stat Methodol* 1996:267-88.

17. Meinshausen N, Bühlmann P. Stability selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2010;72(4):417-73.

18. Hopkins WG, Marshall SW, Batterham AM, et al. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc* 2009;41(1):3-13. doi: 10.1249/MSS.0b013e31818cb278

19. Hopkins WG. A scale of magnitudes for effect statistics. *A new view of statistics* 2002;502

20. Data structures for statistical computing in Python. Proceedings of the 9th Python in Science Conference; 2010.

21. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011;12(Oct):2825-30.

22. Nédélec M, McCall A, Carling C, et al. Recovery in soccer: part I - post-match fatigue and time course of recovery. *Sports Med* 2012;42(12):997-1015. doi: 10.2165/11635270-000000000-00000

23. Lindstedt SL, LaStayo PC, Reich TE. When active muscles lengthen: properties and consequences of eccentric contractions. *News Physiol Sci* 2001;16(6):256-61.

24. Coutts AJ, Kempton T, Sullivan C, et al. Metabolic power and energetic costs of professional Australian Football match-play. *J Sci Med Sport* 2015;18(2):219-24. doi: 10.1016/j.jsams.2014.02.003

25. Brito J, Hertzog M, Nassis GP. Do match-related contextual variables influence training load in highly trained soccer players? *J Strength Cond Res* 2016;30(2):393-9. doi: 10.1519/JSC.0000000000001113

26. McLaren SJ, Smith A, Spears IR, et al. A detailed quantification of differential ratings of perceived exertion during team-sport training. *J Sci Med Sport* 2017;20(3):290-5. doi: 10.1016/j.jsams.2016.06.011

27. Weston M, Siegler J, Bahnert A, et al. The application of differential ratings of perceived exertion to Australian football league matches. *J Sci Med Sport* 2015;18(6):704-8. doi: 10.1016/j.jsams.2014.09.001

28. Arcos AL, Yanci J, Mendiguchia J, et al. Rating of muscular and respiratory perceived exertion in professional soccer players. *J Strength Cond Res* 2014;28(11):3280-8. doi: 10.1519/JSC.0000000000000540

29. Vanrenterghem J, Nedergaard NJ, Robinson MA, et al. Training load monitoring in team sports: a novel framework separating physiological and biomechanical load-adaptation pathways. *Sports Med* 2017:1-8. doi: 10.1007/s40279-017-0714-2 [published Online First: 10 Mar 2017]

30. Gallo TF, Cormack SJ, Gabbett TJ, et al. Pre-training perceived wellness impacts training output in Australian football players. *J Sports Sci* 2016;34(15):1445-51. doi: 10.1080/02640414.2015.1119295

31. Saw AE, Main LC, Gastin PB. Monitoring the athlete training response: subjective self-reported measures trump commonly used objective measures: a systematic review. *Br J Sports Med* 2016;50(5):281-91. doi: 10.1136/bjsports-2015-094758

32. Akubat I, Barrett S, Abt G. Integrating the internal and external training loads in soccer. *Int J Sports Physiol Perform* 2014;9(3):457-62. doi: 10.1123/ijspp.2012-0347

33. Buchheit M, Cholley Y, Lambert P. Psychometric and physiological responses to a pre-season competitive camp in the heat with a 6-hour time difference in elite soccer players. *Int J Sports Physiol Perform* 2015;11(2):176-81. doi: 10.1123/ijspp.2015-0135

34. Spencer M, Lawrence S, Rechichi C, et al. Time-motion analysis of elite field hockey, with special reference to repeated-sprint activity. *J Sports Sci* 2004;22(9):843-50. doi: 10.1080/02640410410001716715

35. Austin DJ, Gabbett TJ, Jenkins DJ. Repeated high-intensity exercise in a professional rugby league. *J Strength Cond Res* 2011;25(7):1898-904. doi: 10.1519/JSC.0b013e3181e83a5b

Table 1: Set ELIs

Category (# ELIs)	**Definition**
Duration (1 ELI)	This ELI defines the duration of the training session.
Distance (17 ELIs)	These ELIs capture the total distance covered, distances covered in different speed zones, and percentages of distances covered at different speeds. The different speed zones considered are: 0-1 km.h^{-1}, 1-7 km.h^{-1}, 7-12 km.h^{-1}, 12-15 km.h^{-1}, 15-20 km.h^{-1}, 20-25 km.h^{-1}, >25 km.h^{-1}.
Speed (8 ELIs)	This group contains ELIs that describe the distance covered per minute and the number of efforts in different speed zones.
Acceleration and deceleration (18 ELIs)	These ELIs capture the accelerations and decelerations, as well as the accelerating and decelerating distance. The ELIs regarding accelerating and decelerating efforts and distance are divided into different zones based upon magnitude (0-1 m.s^{-2}, 1-2 m.s^{-2}, 2-3.5 m.s^{-2} and >3.5 m.s^{-2}).
PlayerLoad (10 ELIs)	This category consists of ELIs based on measures of PlayerLoad. PlayerLoad 3D is calculated based on the changes in accelerations of a player in the X, Y and Z axis. Also, PlayerLoad per meter (i.e, PlayerLoad 3D per total distance covered) and the PlayerLoad per minute are included. Furthermore, it includes PlayerLoad 1D (i.e., PlayerLoad values per axis).
RHIE (13 ELIs)	An RHIE bout was defined as three or more sprints, high-magnitude accelerations or a combination of both within 21 seconds (modified from Spencer et al[34] and Austin et al[35]). This category included measures based on RHIE such as RHIE bout recovery, RHIE duration, RHIE per bout, and RHIE total bouts.

Abbreviations: #, number of; ELI, external load indicator; RHIE, repeated high-intensity effort.

Table 2: Machine learning group models and baseline constructed on season 1 and evaluated on season 2: MAEs, 90% CIs, % diff vs LASSO, and effect sizes of MAEs vs baseline

Method	**Aggregation**	**MAE (90% CI)**	**% diff vs LASSO**	**d**	**Effect size**
ANN	Group	1.09 (1.07 – 1.11)	26.6%	0.06	trivial
LASSO	Group	0.80 (0.78 – 0.82)		0.44	small
Baseline	Group	1.14 (1.12 – 1.16)	29.8%		

Abbreviations: % diff, percentage difference; ANN, artificial neural networks; CI, confidence interval; d, standardized difference; LASSO, least absolute shrinkage and selection operator; MAE, mean absolute error; vs, versus.

Table 3: Overview of ELIs and importance score selected by the LASSO group model

ELI	Importance score	Definition
Acceleration zone 4 efforts	0.515	Number of acceleration efforts above 3.5 $m.s^{-2}$
RHIE per bout – mean	0.513	Average of repeated high-intensity efforts per bout of 21 seconds
Deceleration zone 3 distance	0.510	Decelerating distance between -3.5 and -2 $m.s^{-2}$
Velocity zone 5 distance	0.507	Distance covered between 15-20 $km.h^{-1}$
Acceleration zone 3 efforts	0.507	Number of acceleration efforts between 2 and 3.5 $m.s^{-2}$
PlayerLoad	0.487	Accumulated PlayerLoad measured by accelerometry
Velocity zone 4 distance	0.487	Distance covered between 12-15 $km.h^{-1}$
Minutes	0.471	Training duration
Deceleration zone 4 distance	0.466	Decelerating distance below -3.5 $m.s^{-2}$
PlayerLoad 1D side	0.458	Accumulated PlayerLoad for sideways movements (or medio-lateral axis) measured by accelerometry
Velocity zone 6 efforts	0.428	Efforts between 20-25 $km.h^{-1}$
PlayerLoad 2D	0.384	Accumulated PlayerLoad with exclusion of up- and downwards movements (or longitudinal axis) measured by accelerometry

Abbreviations: ELI, external load indicator; LASSO, least absolute shrinkage and selection operator; RHIE, repeated high-intensity effort.

Table 4: Machine learning models and baseline for season 1 and season 2: MAEs, 90% CIs, % diff vs LASSO, and effect sizes of MAEs vs baseline

Method	**Aggregation**	**Season 1**				**Season 2**			
		MAE (90% CI)	**% diff vs LASSO**	**d**	**Effect size**	**MAE (90% CI)**	**% diff vs LASSO**	**d**	**Effect size**
ANN									
	Individual	0.84 (0.82 – 0.86)	3.6%	0.21	small	0.85 (0.83 – 0.87)	0%	0.33	small
	Group	0.81 (0.79 – 0.83)	2.5%	0.26	small	0.83 (0.81 – 0.85)	-2.4%	0.34	small
LASSO									
	Individual	0.81 (0.76 – 0.86)		0.26	small	0.85 (0.80 – 0.90)		0.33	small
	Group	0.79 (0.75 – 0.83)		0.30	small	0.85 (0.80 – 0.90)		0.33	small
Baseline	Group	0.99 (0.94 – 1.04)	20.2%			1.11 (1.05 – 1.17)	23.4%		

Abbreviations: % diff, percentage difference; ANN, artificial neural networks; CI, confidence interval; d, standardized difference; LASSO, least absolute shrinkage and selection operator; MAE, mean absolute error; vs, versus.