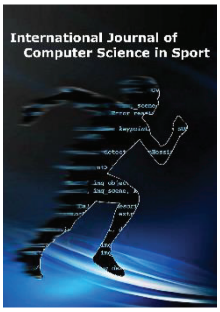
International Journal of Computer Science in Sport

Volume 21, Issue 2, 2022

Journal homepage: http://iacss.org/index.php?id=30

sciendo

DOI: 10.2478/ijcss-2022-0007



Analysis of Relationship between Training Load and Recovery Status in Adult Soccer Players: a Machine Learning Approach

Mandorino, M.[1,2], *Figueiredo, A.J.*[3], *Cima, G.*[4], *Tessitore, A.*[1]

[1]*Department of Movement, Human and Health Sciences, University of Rome "Foro Italico" Rome, Italy*

[2]*Performance and Analytics Department, Parma Calcio 1913, 43121, Parma, Italy*

[3]*University of Coimbra, Research Center for Sport and Physical Activity, Faculty of Sport Science and Physical Education, Coimbra, Portugal*

[4]*Computer, Control and Management Engineering Department, Sapienza University of Rome, Rome, Italy.*

Abstract

Periods of intensified training may increase athletes' fatigue and impair their recovery status. Therefore, understanding internal and external load markers-related to fatigue is crucial to optimize their weekly training loads. The current investigation aimed to adopt machine learning (ML) techniques to understand the impact of training load parameters on the recovery status of athletes. Twenty-six adult soccer players were monitored for six months, during which internal and external load parameters were daily collected. Players' recovery status was assessed through the 10-point total quality recovery (TQR) scale. Then, different ML algorithms were employed to predict players' recovery status in the subsequent training session (S-TQR). The goodness of the models was evaluated through the root mean squared error (RMSE), mean absolute error (MAE), and Pearson's Correlation Coefficient (r). Random forest regression model produced the best performance (RMSE=1.32, MAE=1.04, r = 0.52). TQR, age of players, total decelerations, average speed, and S-RPE recorded in the previous training were recognized by the model as the most relevant features. Thus, ML techniques may help coaches and physical trainers to identify those factors connected to players' recovery status and, consequently, driving them toward a correct management of the weekly training loads.

KEYWORDS: SOCCER, RECOVERY, MACHINE LEARNING, PREDICTION, PERFORMANCE

Introduction

During training and competitive match-play, soccer players perform high-intensity actions such as sprints, jumps, accelerations, and change of directions (Brownstein et al., 2017). Consequently, participation in soccer may lead to experience acute fatigue (immediately after the match), or residual fatigue (still evident up to 72 h post-match), which are conditions characterized by neuromuscular alterations, perturbations of the biochemical milieu, and the psychometric state (Carling et al., 2018; Hader et al., 2019). It was recognized that a poor or insufficient recovery during training sessions, and the resulting fatigue, by affecting physical performance and technical activity, could increase the athletes' risk of injury (Clemente, Figueiredo, Martins, Mendes, & Wong, 2016; Mandorino, Figueiredo, Cima, & Tessitore, 2021; Mandorino, Figueiredo, Cima, & Tessitore, 2022; Mandorino, Figueiredo, Condello, & Tessitore, 2022).

For this reason, monitoring athletes is considered essential for understanding their responses to training and match load, to maximize performance and ensure their competition readiness (Halson, 2014; Lacome et al., 2018). An athlete's training load can be quantified using external (e.g., total distance, velocity, accelerations) and internal parameters (e.g., heart rate, lactate, rate of perceived exertion (RPE)). External training load has been defined as the work completed by an athlete, and assessed independently by his or her internal characteristics. Instead, the internal training load represents the relative physiological and psychological stress imposed on the athlete (Halson, 2014; Impellizzeri et al., 2004). The combination of external and internal parameters is considered the optimal condition for quantifying the overall training load (Halson, 2014). Differently, in order to quantify fatigue and recovery status after training or matches, the use of athletes' self-reported measures (e.g., total quality recovery scale (TQR) or questionnaires (e.g., subjective wellbeing questionnaire) is recommended (Saw et al., 2016).

Assuming that changes in recovery status and the consequent fatigued condition could affect physical performance and increase the risk of injury, it is crucial to analyze the influence of training load on players' recovery status. Thorpe et al. (2015, 2017) observed that changes in total high-intensity running distance (THIR) were correlated with fluctuations of fatigue. Similarly, Fessi et al. (2016) identified a positive correlation between the daily training load and rating of sleep, stress, fatigue, and muscle soreness. Instead, Selmi et al. (2020) observed that during intensified training periods characterized by high RPE, monotony, and strain values, significant lower TQR scores were recorded in professional soccer players. In addition, low recovery values could, in turn, influence the subsequent training output. Indeed, Malone et al. (2018) found that a reduction in individual well-being scores significantly impacted on players' running performance during the subsequent training session. In summary, a high training load impairs athletes' recovery status, negatively affecting their physical performance during the subsequent training or match.

Therefore, daily monitoring of players' recovery status would detect early signs of tiredness and maximize their training and match performance (Selmi et al., 2018). In addition, knowing in advance changes of the players' recovery status may allow coaches and physical trainers to accurately program and adapt training loads in relation to their individual needs. If previous studies deeply investigated the relationship between training load and recovery (Sawczuk et al., 2018; Vescovi et al., 2019), to the best of our knowledge, there are few studies that tried to predict players' recovery status using internal and external load data. In order to achieve this goal, machine learning (ML) techniques could be useful for this type of analyses as they (1) allow to build predictive models, (2) manage a large set of variables, (3) identify non-linear relationship between training load and players' recovery status, (4) identify most relevant features (Bunker & Thabtah, 2019; Cai et al., 2018; Camacho et al., 2018; Ray, 2019).

Thus, the purpose of the current study was to predict the recovery status of adult soccer players using external and internal load data by adopting ML techniques. The scope was to identify a tool that allows coaches and physical trainers to know in advance the players' recovery status and, consequently, to optimize the management of training loads in subsequent training sessions. In addition, the current study aimed to identify the most relevant parameters that affected the recovery status and analyze the impact of itsvariation on players' performance.

This article represents an extended version of (Mandorino, Figueiredo, Cima, et al., 2022b), which focused on the performance analysis of different ML models. The current study expands (Mandorino, Figueiredo, Cima, et al., 2022b) in several ways. The main added contributions are the following:

1. A more detailed description of data collection was provided. Furthermore, all the features used in ML models were described.
2. Feature engineering, data pre-processing, validation, and model evaluation were explained in more detail.
3. Bland-Altman analysis was conducted to evaluate the agreement between ML predictions and real values.
4. The relationship between variation in recovery status and players' performance was analyzed.
5. Practical applications were provided to support coaches and physical trainers in managing the weekly training load.

Methods

Participants

Twenty-six adult non-professional male soccer players (mean ± SD age: 21.3 ± 4.3 years, height: 178.2 ± 7.2 cm, body mass: 73.0 ± 7.0 kg) were monitored for six months, from September to February, during the 2019/2020 soccer season. Due to the Covid pandemic restrictions, the soccer season was interrupted and, consequently, also our data collection. All movement players were included in this study: central backs (n=4), fullbacks (n=6), midfielders (n=8), forwards (n=8). Instead, goalkeepers were excluded. All players trained four days per week (Tuesday, Wednesday, Thursday, and Friday) and participated in an official match during the weekend (Sunday). The week's days were categorized as days before the match day, i.e., MD minus (MD-5, MD-4, MD-3, MD-2, MD). The study was conducted in accordance with the Declaration of Helsinki, and it was approved by the local research ethics committee.

Training/Match Load and recovery status

Players' physical activity was recorded during 62 training sessions and 17 matches. A total of 1168 individual sessions corresponding to the 79 collective sessions was then monitored. External training/match load was obtained using Johan GPS (JOHAN Sports, Noordwijk, Netherlands) consisting of a GPS sensor based on navigation technology from the European Space Agency (10 Hz, including EGNOS correction), accelerometer, gyroscope, and magnetometer (100 Hz, 3 axes, ±16 g), which validity and reliability was assessed by Nikolaidis et al. (2018). The GPS devices were placed between the players' scapulae through a tight vest. Seventeen workload variables, including kinematic (e.g., total distance, distance covered at specific velocity) and mechanical variables (e.g., Player Load, number of accelerations/decelerations above specific thresholds), were extracted from the GPS data (Rossi et al., 2018). Moreover, the session-RPE method (S-RPE) (Foster et al., 1995) was used to quantify the internal training/match load. Athlete's session-RPE were collected about 30 min after the end of each training session (Impellizzeri et al., 2004). The S-RPE scores were obtained by multiplying the rate of perceived exertion (RPE) value, quantified through the CR-10 Borg's scale modified by Foster et al. (1995), by the duration of each training or match for every single player. The perceived recovery status of players was quantified using the 10-point total quality recovery scale (TQR). Based on their personal psychophysical cues (e.g., mood states, muscle soreness), the atheletes quantified their recovery status fifteen minutes before the warm-up of the training session or before the match. This method was previously introduced by Telenttä and Hassmén (1998), and it was already successfully used to quantify the recovery status in soccer players (Gjaka et al., 2016) and basketball players (Sansone et al., 2020). TQR was included, as we will see later, in the list of features to predict the TQR score of the subsequent training session (S-TQR).

Feature Engineering

In addition to internal and external training/match load data collected with GPS and S-RPE method, other features were added. Personal players' information (age and role of play) was inserted in the ML models. To quantify the weekly external load (WEL), the rolling sum with a span of 7 days was calculated for each of the seventeen external load parameters. To quantify the weekly internal load (WIL), the rolling sum of S-RPE with a span of 7 days was calculated (Rossi et al., 2019). Moreover, the cumulative loads for a period of 2, 3, and 4 weeks (WIL2, WIL3, WIL4, respectively) were calculated. In addition to the internal load parameters, the acute:chronic workload ratio (ACWR) was determined by dividing the weekly workload (acute load) by the average weekly workload over the previous 4 weeks (chronic load) (Gabbett, 2016; Malone et al., 2017). Particularly, ACWR was daily calculated by dividing the rolling 7-day workload by the chronic rolling 28-day workload (Murray et al., 2017). Considering that the management of the weekly load changes in relation to distance from the previous and next match, information regarding the day of the week was added to the list of features.

Moreover, since the perception of fatigue could change as the season goes by, the month during which the training session has been performed was added (Rossi et al., 2019). Finally, the number of recovery days from the previous training was also inserted. Therefore, a total of 50 features was considered in the current study. The features are presented and summarized in Table 1.

Data Analysis

In the current study, multiple ML classifiers were built using a training dataset where each example describes the athlete's training session. Each training session consists of a vector of features representing the players' recent workload. For this reason, the days without training were not represented by any vector. The different ML classifiers were constructed to predict the players' recovery status in the subsequent training session. Therefore, S-TQR was employed in the ML models as label.

Table 1. Summary of external load, internal load, recovery features, and players' personal information

Features		
Personal	Age and role	
Contextual	Month, day of the week, and number of days of recovery	
Internal Load Parameters		
Daily	RPE, S-RPE	
Weekly	WIL, WIL2, WIL3, WIL4	
ACWR	Acute (span = 7 days) to chronic (span = 28 days) workload ratio	
External Load Parameters		
Daily	*Kinematic features*	Duration of training, total distance, walk distance (< 7 km/h), jog distance (≥ 7 km/h and < 14 km/h), run distance (≥14 km/h and < 20 km/h), sprint distance ≥ 20 km/h and < 25 km/h), high sprint distance (≥ 25 km/h), max speed, average speed, number of sprints, number of high intensity sprints, number of repeated sprints
	Mechanical features	Playerload 2D, playerload 3D, total accelerations (> 2 m/s2), total decelerations (> 2 m/s2), high accelerations (> 3 m/s2), high decelerations (> 3/ms2)
Weekly	The rolling sum with a span of 7 days was calculated for all external load parameters	
Perceived Recovery Status		
Daily	TQR	

ML Algorithms Selection

Different algorithms were selected to test their ability to predict the recovery status of our players. The set of features was inserted in the model as predictors and employed to forecast players' recovery status in the subsequent training session (S-TQR). The target variable was treated as a continuous variable. Although the target variable is ordinal, the Likert scale with five or more categories can be treated as continuous without affecting the statistical analysis (Johnson & Creech, 1983; Norman, 2010; Rhemtulla et al., 2012; Robitzsch, 2020). Considering the scarcity of ML algorithms for ordinal classification (Frank & Hall, 2001), this approach allows more flexibility and relies on the conventional ML models.

The following classifiers were built: linear regression (LR), support vector regression (SVR), decision tree regression (DT), random forest regression (RF). Describing the underlying mathematical functions of the models is outside the scope of this paper. However, LR was selected for its ability to understand the linear relationship between input and output numerical variables. Differently, SVR, DT, and RF were chosen for their ability to model complex and non-linear interactions inside high-dimensional data (Kensert et al., 2018).

Data pre-processing

Standard pre-processing techniques were used to optimize the performance of the different models. Firstly, a data cleaning process was applied. The days in which, for any reason, players did not wear the GPS vest were excluded from the analysis. Out of 1205 observations, 37 (3.1%) were excluded. The advantage of removing data considered "missing completely at random" is that the analysis remains unbiased (Kang, 2013). In addition, missing internal load and recovery data were replaced by the mean value of the player's corresponding parameter. All the features were standardized adopting the Standard Scaler (SS). SS is a scaling method that normalizes each feature by subtracting the mean and then scaling to unit variance, which means dividing by the standard deviation (Ferreira et al., 2019). Normalization ensures that all the features fair contributes to the learning process (Singh & Singh, 2020).

Experiments

First, a feature selection process was performed on 20% of the dataset. Recursive Feature Elimination (RFE) was performed to identify and remove correlated features that could increase overfitting risk (Rossi et al., 2018). The RFE feature selection algorithm was set to identify the fifteen most relevant features in predicting the target variable. After this first step, the classifiers were validated on the remaining 80% of the dataset. Nested cross-validation was performed with stratified 3-fold-cross-validation in the outer layer and 3 fold cross-validated grid search in the inner layer (Murugesan et al., 2018). The inner layer performed the grid search for model (hyperparameter) selection within each classifier. The root-mean-square error (RMSE) was chosen as the performance metric. Nested cross-validation holds out test data from training data and allows to obtain an unbiased estimation of the real-world performance (Cawley & Talbot, 2010; Murugesan et al., 2018). All analyses were performed using Anaconda and Python libraries.

Model Evaluation

The goodness of the classifiers was evaluated adopting the mean absolute error (MAE), and root mean squared error (RMSE). Low values of MAE and RMSE indicate model prediction ability. The model's goodness was also assessed by analyzing the relationship between S-TQR observed and predicted using the Pearson correlation coefficient. Particularly, the average Pearson correlation was calculated across the cross-validation runs. The Pearson correlation coefficient (r) can range from -1 (negative correlation) to +1 (positive correlation). In addition, as proposed in Rossi et al. (2019), Bland-Altman analysis was employed to evaluate the agreement between the real S-TQR and the predicted value. This analysis allows to quantify the bias (i.e., mean difference between S-TQR observed and predicted), and systematic error (i.e., the relationship between mean and difference in S-TQR predicted and observed) of the classifiers. Moreover, the performance of the classifiers was compared with two baselines: baseline B_1 generated predictions by respecting the training set's class distribution; baseline B_2 always predicted the most frequent label in the training set.

Analysis of relationship between variation in recovery status and players' performance

The model which produced the best performance was selected for further analysis. Considering the predictions made by the model, four conditions were identified:

- **Severe Lowering of the recovery status (SLRS):** the classifier predicted a decrease in the player's recovery status of more than 20%;
- **Lowering of the recovery status (LRS):** the classifier predicted a decrease in the player's recovery status between 1% and 20%;
- **Increase of the recovery status (IRS):** the classifier predicted an increase in player's recovery status between 1% and 20%;
- **Strong Increase of the recovery status (SIRS):** the classifier predicted an increase in player's recovery status higher than 20%.

One-way ANOVA was employed to compare players' percentage performance variation (RPE, High-intensity sprint distance, total acceleration, and total deceleration) in the following training session considering the classifier's prediction (SLRS, LRS, IRS, and SIRS). One-way ANOVA was performed using the Statistical Package for the Social Science, version 25.0 (SPSS Inc., Chicago, IL, USA). The level of statistical significance was set at $p < 0.05$.

Analysis of relationship between recovery status and players' age

Repeated measure correlation (rmcorr) was performed to analyze the relationship between recovery status, assessed throughout the season, and players' age, adjusting for the visit effect (training sessions) (Bakdash & Marusich, 2017; Shan et al., 2020). The analysis was conducted using Anaconda and Python libraries. The level of statistical significance was set at $p < 0.05$.

Results

Analysis of internal (S-RPE) and external (total distance) load distribution during the week was presented in Figure 1. Among the different classifiers, RF produced the best performance (MAE $=1.043$, RMSE $= 1.321$, and $r = 0.521$). The performance of the other classifier was presented in Figure 2. Producing the best performance, RF was involved in further analysis. After RFE analysis, only 15 features were selected.

Moreover, feature importance analysis was presented in Figure 3. TQR of the previous day, age of the players, total decelerations, average speed, and S-RPE were recognized as the five most important features during the cross-validation. Bland-Altman analysis, presented in Figure 4, was performed to quantify the agreement between the real S-TQR and the predicted value. A low bias between the two measures was observed (0.014 ± 1.36 arbitrary units (AU)).

Moreover, a simple linear regression was fit between the mean and the difference of real S-TQR and the predicted value. The slope of the regression line was interpreted as the systematic error (Rossi et al., 2019). A negative slope was found ($m = -0.91$). The negative slope highlights that, for low values, the model tends to overestimate the S-TQR. Conversely, as soon as the values increase, the trend changes, and the model underestimates the S-TQR.

After physical performance analysis, One-way ANOVA revealed a significant difference between the four conditions previously identified (SLRS, LRS, IRS, SIRS; $p<0.05$). The SIRS condition produced a significant percentage increase in physical performance (High-intensity sprint distance, total accelerations, total decelerations) compared to the SLRS one ($p < 0.05$). Moreover, significant higher RPE values were found in the SIRS condition compared to the LRS ($p < 0.05$) and SLRS ones ($p < 0.01$). The results were summarized in Figure 5.

Rmcorr analysis revealed a significant negative correlation between S-TQR and players' age but a weak association ($r = -0.21$, $p < 0.01$; Figure 6).

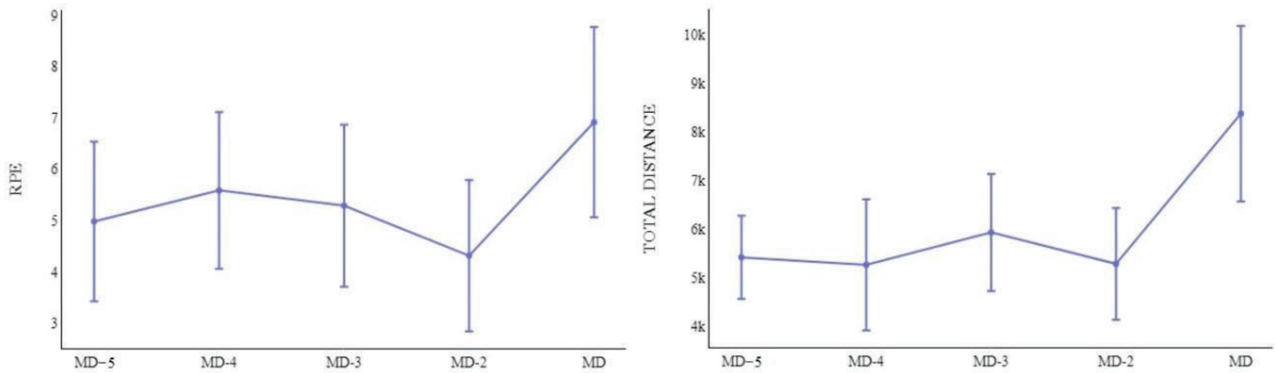


Fig. 1. Analysis of RPE and total distance distribution during the week. Values are reported as mean ± SD. SD = standard deviation. MD = match day

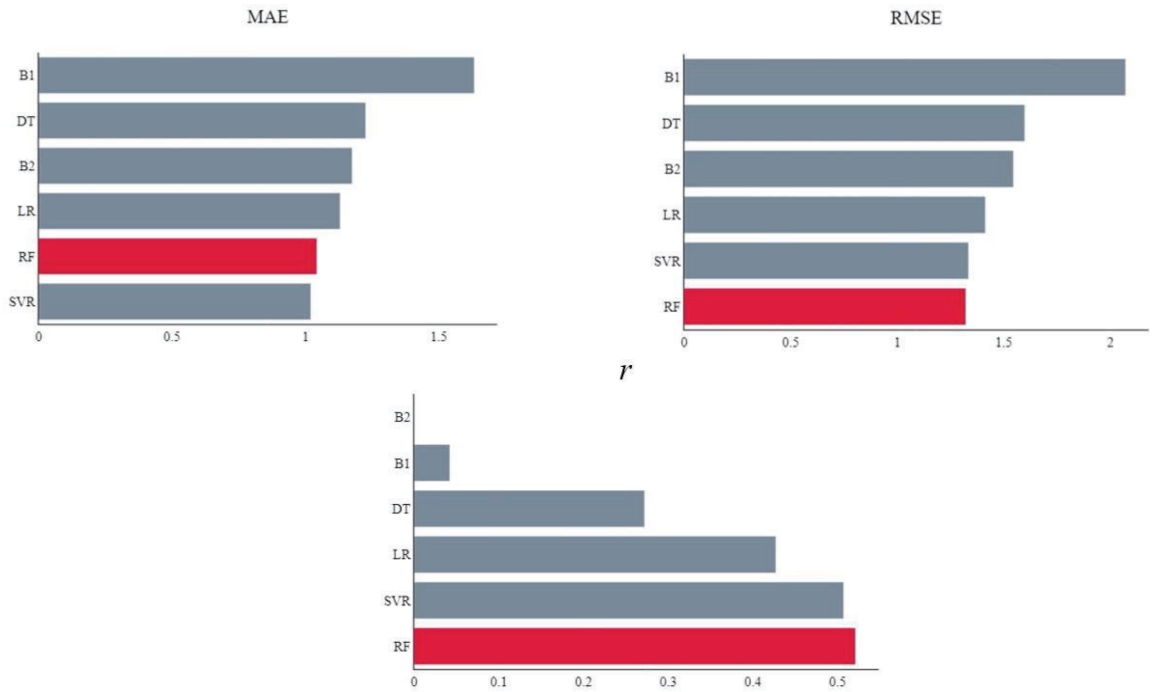


Fig. 2. Performance of the different classifiers. MAE = mean absolute error. RMSE = root mean squared error. R = Pearson correlation coefficient. B1 = Baseline 1. B2 = Baseline 2. LR = linear regression. RF = random forest regression. DT = decision tree regression. SVR = support vector regression

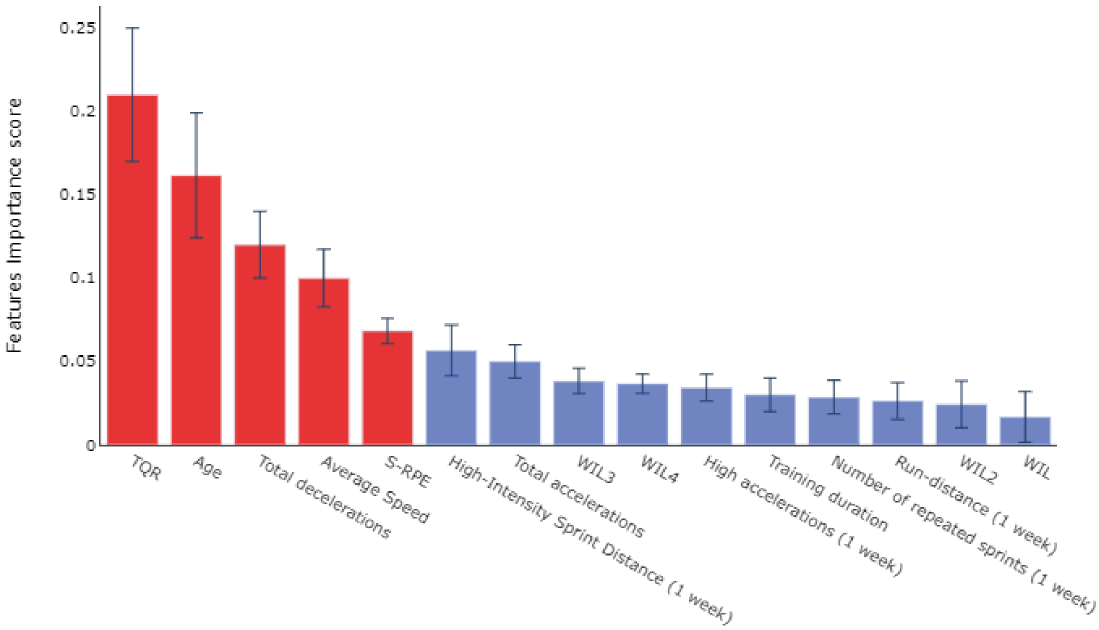


Fig. 3. Feature importance analysis.

Discussion

The main purpose of the current study was to exploit ML techniques to predict soccer players' recovery status in the following training session. Being able to predict in advance athletes' recovery status would allow to promptly detect signs of tiredness and accurately program and adapt their training loads. A similar approach was previously developed by De Beéck et al. et al. et al. (2019), who differently assessed perceived player wellness through a questionnaire consisting of five different questions about fatigue, sleep quality, general muscle soreness, stress levels, and mood. In the current study, recovery status was assessed through TQR scale, which, compared to the previous approach, has the advantage of quantifying players' recovery perception with a single value. Different external load and internal load data and different ML algorithms were employed for the purpose of this study. Among the various ML techniques selected, RF produced the best performance (MAE =1.043, RMSE = 1.321, r = 0.521; Figure 2) outperforming the baselines (B_1 and B_2). RF is a three-based ensemble method, and ensemble-based methods generally perform better than the individual learners that construct them (Ahmad et al., 2018). This aspect is related to their ability to combine multiple ML techniques into one predictive model reducing variance, bias and increasing predictive ability (Zhang et al., 2012).

Being the most performing algorithm, RF was involved in further analysis. Among the features selected after RFE analysis, TQR of the previous day, age of the players, total decelerations, average speed, and S-RPE (Figure 3) were identified as the most important features in predicting the target variable. In accordance with De Beéck et al. (2019), these results suggest that the combination of internal and external load, together with preceding perceived recovery status, represents the optimal condition to predict players' recovery in the following training session (S-TQR). In particular, total decelerations were recognized as the most relevant among the external load parameters. Similarly, Jaspers et al. (2018) found that decelerations were the load that most affected players' RPE. Decelerating efforts produce eccentric contractions mainly associated with muscle damage (Nédélec et al., 2012). Therefore, a large number of decelerations may increase athletes' perception of effort, and consequently, they may increase fatigue and alter recovery status in the following training session. Coaches and physical trainers should be aware that a high training load could increase players' neuromuscular fatigue, and they should recognize the type of load that has the most significant influence on it.

In addition to training load data, it must be considered personal information of the players. Indeed, age was identified as the second most relevant feature. More specifically, as evidenced in Figure 6, the average recovery status tends to decrease over the years. To make the graph clearer, the average recovery status was presented in relation to players' age and the day of the week. The relationship between S-TQR and age of the player was investigated adopting the rmcorr technique. The analysis was conducted while controlling for the visit effect (training sessions) (Bakdash & Marusich, 2017; Shan et al., 2020). It showed a significant negative association between recovery status and the age of the players. Although the subjects involved in the present study were young (age: 21.3 ± 4.3 years), it is well accepted that aging processes could affect athletes' perception of recovery (Fell & Williams, 2008). This result could be explained by the fact that aging skeletal muscle experiences greater fatigue or damage, and consequently slower rate of repair and recovery (Fell & Williams, 2008). However, we must emphasize that the analysis revealed only a weak correlation, and this could be explained by the small sample size and the narrow age range. In general, it is necessary to consider that the impact of training load could change in relation to individual aspects as the age of the players. Indeed, as reported in previous studies, the individual characteristics of the athletes (e.g., aerobic fitness level, age, body composition) could significantly impact the stress placed on the body, and consequently, increase predisposition to injury (Impellizzeri et al., 2005; Jones et al., 2017; Kalkhoven et al., 2021).

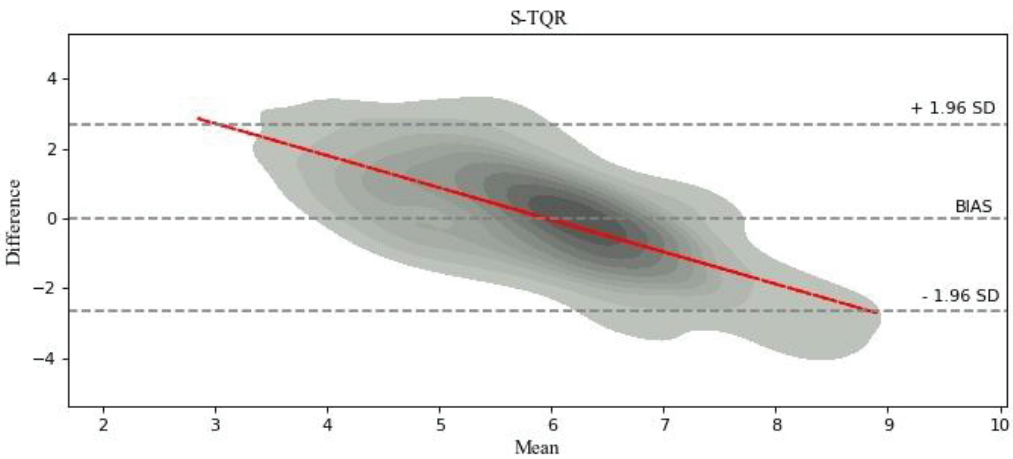


Fig. 4. Bland-Altman Analysis. Analysis of relationship between the mean and the difference among the observed and predicted S-TQR values. The mean of the difference (i.e., bias) and the 1.96 standard deviation to the mean of the difference (i.e., confidence interval) were reported. The red line represents the relationship between mean and difference (i.e., systematic error).

Although the RF algorithm was able to predict S-TQR values with a low bias (0.014 ± 1.36 arbitrary units (AU)), as emerged from Bland-Altman analysis (Figure 4), the accuracy of the model differed as the S-TQR changed. Particularly, the model overestimated the recovery status at low values of S-TQR and underestimated it at high values. The same issue was addressed in Rossi et al. (2019), who explained the problem through the impact of psychological aspects that cannot be considered inside ML techniques. However, the problem could also be explained by considering the nature of the training set. Indeed, the low and high S-TQR values are less represented. Consequently, the model may exhibit less learning ability for these values. Future studies could adopt sampling strategies for regression tasks (SMOTE for regression) to limit the effect of class imbalance.

To understand the effect of recovery status, fluctuations in performance (RPE, high-intensity sprint distance, total acceleration, total deceleration) were analyzed in relation to the variation of the recovery status predicted by the RF model. Four different classes describing a reduction (SLRS, LRS) or increase (IRS, SIRS) in recovery status were identified. It was observed that when the model predicted a severe lowering of players' recovery status, a significant decrease in players' performance was also registered (Figure 5). In line with previous studies, poor recovery and fatigue may influence technical activity and reduce physical intensity during the training session (Clemente et al., 2016; Selmi et al., 2018). Malone et al. (2018) found that a reduction in wellbeing score was associated with running performance impairment during the subsequent training session. As further proof of this phenomenon, Tessitore et al. (2007) observed that the adoption of different recovery strategies increased soccer players' anaerobic performance during preseason. Therefore, we can reasonably argue that estimating players' recovery status in advance would allow to program, adapt training loads, and implement individualized recovery strategies.

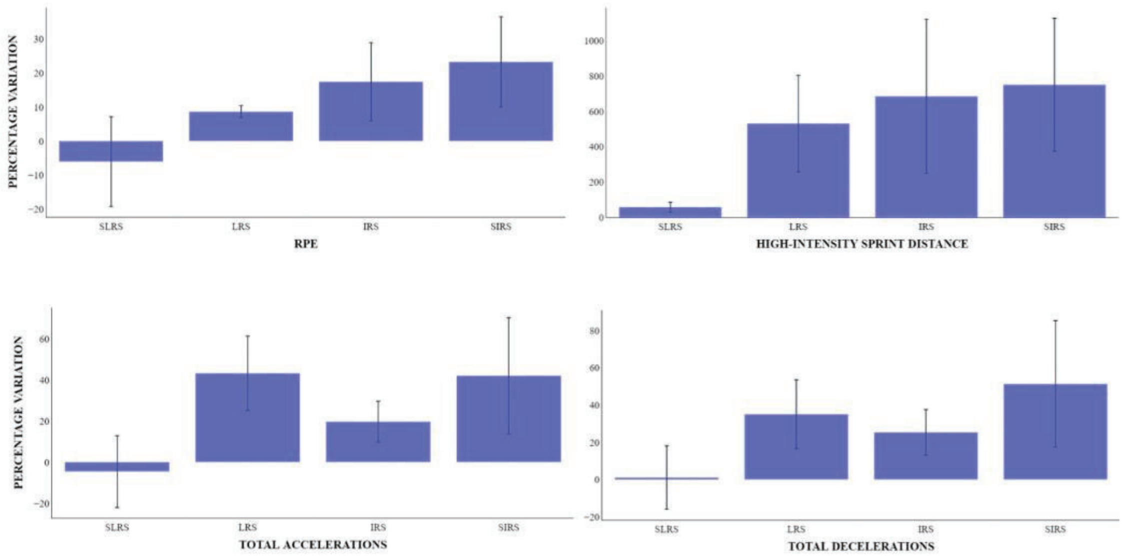


Fig. 5. Physical Performance Analysis. SLRS = Severe Lowering of the recovery status**.** LRS = Lowering of the recovery status**.** IRS = Increase of the recovery status**.** SIRS = Strong Increase of the recovery status

Limitations of the study

Considering the limitations of the current research, future studies should replicate this design by increasing the sample size and over consecutive seasons. Moreover, the use of additional parameters (e.g., heart rate variability, blood markers) could increase the predictive ability of ML models. Moreover, the study is limited to non-professional soccer players; therefore, future studies could replicate this design involving other populations (e.g., professional soccer players, rugby players). Finally, the GPS system adopted in the current study did not allow to collect the horizontal dilution of position (HDOP), which reflects the signal's accuracy; therefore, it was not possible to check the quality of the signal.

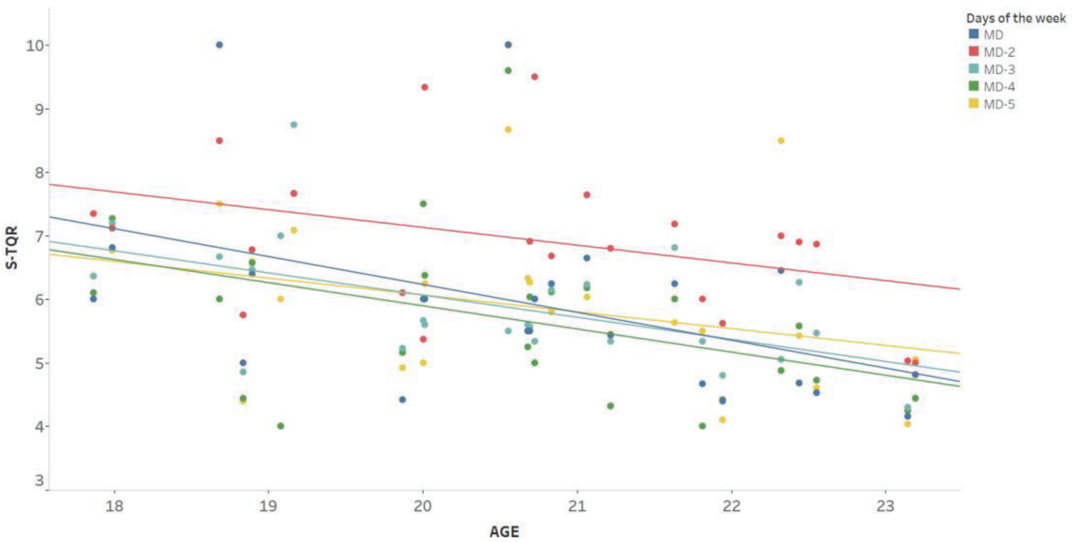


Fig. 6. Analysis of the average recovery status in relation to the age of the soccer players. The average of players' recovery status was grouped in relation to the days of the week.

Practical applications

A reduction in athletes' recovery status could increase the risk of injury (Brink et al., 2010) and impair physical performance in the following training session (Malone et al., 2018). Therefore, the use of ML techniques could help coaches and physical trainers predict the players' recovery status. Knowing in advance players' recovery status would allow to:

- Optimize the training load: it might be useful to reduce the training load to players who register a severe decrease in recovery status. Considering the results reported in the current study, training could be individualized acting on the parameters that most overload players (e.g., total decelerations);
- Implement personalized recovery strategies: players with a low recovery status may be subjected to personalized recovery strategies before the training (e.g., mobility exercises, massage, and tissue treatments).

Conclusion

Optimizing the recovery status of soccer players is essential to maximize their performance and reduce the risk of injury. The ML approach allowed predicting the S-TQR score and understanding the most influential variables on players' recovery status. The analysis revealed the importance of including internal and external parameters to ensure an optimal monitoring strategy and understand the soccer players' recovery status. The model developed may help coaches and physical trainers in the correct management of the weekly training load.

Acknowledgements

The authors would like to thank the club Atletico Lodigiani, all the players, and the coaching staff (Marco Di Rocco, Marco Ciampi, Marco Lista), for their participation in the study.

References

Ahmad, M. W., Reynolds, J., & Rezgui, Y. (2018). Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees. *Journal of cleaner production*, *203*, 810–821.

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in psychology*, *8*, 456.

Brink, M. S., Visscher, C., Arends, S., Zwerver, J., Post, W. J., & Lemmink, K. A. (2010). Monitoring stress and recovery: New insights for the prevention of injuries and illnesses in elite youth soccer players. *British journal of sports medicine*, *44*(11), 809–815.

Brownstein, C. G., Dent, J. P., Parker, P., Hicks, K. M., Howatson, G., Goodall, S., & Thomas, K. (2017). Etiology and recovery of neuromuscular fatigue following competitive soccer match-play. *Frontiers in physiology*, *8*, 831.

Bunker, R. P., & Thabtah, F. (2019). A machine learning framework for sport result prediction. *Applied computing and informatics*, *15*(1), 27–33.

Cai, J., Luo, J., Wang, S., & Yang, S. (2018). Feature selection in machine learning: A new perspective. *Neurocomputing*, *300*, 70–79.

Camacho, D. M., Collins, K. M., Powers, R. K., Costello, J. C., & Collins, J. J. (2018). Next-generation machine learning for biological networks. *Cell*, *173*(7), 1581–1592.

Carling, C., Lacome, M., McCall, A., Dupont, G., Le Gall, F., Simpson, B., & Buchheit, M. (2018). Monitoring of post-match fatigue in professional soccer: Welcome to the real world. *Sports Medicine*, *48*(12), 2695–2702.

Cawley, G. C., & Talbot, N. L. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *The Journal of Machine Learning Research*, *11*, 2079–2107.

Clemente, F. M., Figueiredo, A. J., Martins, F. M. L., Mendes, R. S., & Wong, D. P. (2016). Physical and technical performances are not associated with tactical prominence in U14 soccer matches. *Research in Sports Medicine*, *24*(4), 352–362.

De Beéck, T. O., Jaspers, A., Brink, M. S., Frencken, W. G., Staes, F., Davis, J. J., & Helsen, W. F. (2019). Predicting future perceived wellness in professional soccer: The role of preceding load and wellness. *International Journal of Sports Physiology and Performance*, *14*(8), 1074–1080.

Fell, J., & Williams, A. D. (2008). The effect of aging on skeletal-muscle recovery from exercise: Possible implications for aging athletes. *Journal of Aging and Physical Activity*, *16*(1), 97–115.

Ferreira, P., Le, D. C., & Zincir-Heywood, N. (2019). Exploring feature normalization and temporal information for machine learning based insider threat detection. *2019 15th International Conference on Network and Service Management (CNSM)*, 1–7.

Fessi, M. S., Nouira, S., Dellal, A., Owen, A., Elloumi, M., & Moalla, W. (2016). Changes of the psychophysical state and feeling of wellness of professional soccer players during pre-season and in-season periods. *Research in Sports Medicine*, *24*(4), 375–386.

Foster, C., Hector, L. L., Welsh, R., Schrager, M., Green, M. A., & Snyder, A. C. (1995). Effects of specific versus cross-training on running performance. *European journal of applied physiology and occupational physiology*, *70*(4), 367–372.

Frank, E., & Hall, M. (2001). A simple approach to ordinal classification. *European conference on machine learning*, 145–156.

Gabbett, T. J. (2016). The training—Injury prevention paradox: Should athletes be training smarter and harder? *British journal of sports medicine*, *50*(5), 273–280.

Gjaka, M., Tschan, H., Francioni, F. M., Tishkuaj, F., & Tessitore, A. (2016). MONITORING OF LOADS AND RECOVERY PERCEIVED DURING WEEKS WITH DIFFERENT SCHEDULE IN YOUNG SOCCER PLAYERS. *Kinesiologia Slovenica*, *22*(1).

Hader, K., Rumpf, M. C., Hertzog, M., Kilduff, L. P., Girard, O., & Silva, J. R. (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. *Sports medicine-open*, *5*(1), 1–19.

Halson, S. L. (2014). Monitoring training load to understand fatigue in athletes. *Sports medicine*, *44*(2), 139–147.

Impellizzeri, F. M., Rampinini, E., Coutts, A. J., Sassi, A., & Marcora, S. M. (2004). Use of RPE-based training load in soccer. *Medicine & Science in sports & exercise*, *36*(6), 1042–1047.

Impellizzeri, F. M., Rampinini, E., & Marcora, S. M. (2005). Physiological assessment of aerobic training in soccer. *Journal of sports sciences*, *23*(6), 583–592.

Jaspers, A., De Beéck, T. O., Brink, M. S., Frencken, W. G., Staes, F., Davis, J. J., & Helsen, W. F. (2018). Relationships between the external and internal training load in professional soccer: What can we learn from machine learning? *International journal of sports physiology and performance*, *13*(5), 625–630.

Johnson, D. R., & Creech, J. C. (1983). Ordinal measures in multiple indicator models: A simulation study of categorization error. *American Sociological Review*, 398–407.

Jones, C. M., Griffiths, P. C., & Mellalieu, S. D. (2017). Training load and fatigue marker associations with injury and illness: A systematic review of longitudinal studies. *Sports medicine*, *47*(5), 943–974.

Kalkhoven, J. T., Watsford, M. L., Coutts, A. J., Edwards, W. B., & Impellizzeri, F. M. (2021). Training load and injury: Causal pathways and future directions. *Sports Medicine*, *51*(6), 1137–1150.

Kang, H. (2013). The prevention and handling of the missing data. *Korean journal of anesthesiology*, *64*(5), 402.

Kensert, A., Alvarsson, J., Norinder, U., & Spjuth, O. (2018). Evaluating parameters for ligand-based modeling with random forest on sparse data sets. *Journal of cheminformatics*, *10*(1), 1–10.

Kenttä, G., & Hassmén, P. (1998). Overtraining and recovery. *Sports medicine*, *26*(1), 1–16.

Lacome, M., Simpson, B., Broad, N., & Buchheit, M. (2018). Monitoring players' readiness using predicted heart-rate responses to soccer drills. *International Journal of Sports Physiology and Performance*, *13*(10), 1273–1280.

Malone, S., Owen, A., Newton, M., Mendes, B., Collins, K. D., & Gabbett, T. J. (2017). The acute: Chonic workload ratio in relation to injury risk in professional soccer. *Journal of science and medicine in sport*, *20*(6), 561–565.

Malone, S., Owen, A., Newton, M., Mendes, B., Tiernan, L., Hughes, B., & Collins, K. (2018). Wellbeing perception and the impact on external training output among elite soccer players. *Journal of science and medicine in sport*, *21*(1), 29–34.

Mandorino, M., Figueiredo, A. J., Cima, G., & Tessitore, A. (2021). A Data Mining Approach to Predict Non-Contact Injuries in Young Soccer Players. *International Journal of Computer Science in Sport*, *20*(2), 147–163.

Mandorino, M., Figueiredo, A. J., Cima, G., & Tessitore, A. (2022a). Predictive Analytic Techniques to Identify Hidden Relationships between Training Load, Fatigue and Muscle Strains in Young Soccer Players. *Sports*, *10*(1), 3.

Mandorino, M., Figueiredo, A. J., Cima, G., & Tessitore, A. (2022b). The Impact of External and Internal Load on Recovery Status of Adult Soccer Players: A Machine Learning Approach. *International Conference on Security, Privacy, and Anonymity in Computation, Communication, and Storage*, 122–125.

Mandorino, M., Figueiredo, A. J., Condello, G., & Tessitore, A. (2022). The influence of maturity on recovery and perceived exertion, and its relationship with illnesses and non-contact injuries in young soccer players. *Biology of Sport*, *39*(4), 839–848.

Murray, N. B., Gabbett, T. J., Townshend, A. D., & Blanch, P. (2017). Calculating acute: Chronic workload ratios using exponentially weighted moving averages provides a more sensitive indicator of injury likelihood than rolling averages. *British Journal of Sports Medicine*, *51*(9), 749–754.

Murugesan, G., Saghafi, B., Davenport, E., Wagner, B., Urban, J., Kelley, M., Jones, D., Powers, A., Whitlow, C., & Stitzel, J. (2018). Single season changes in resting state network power and the connectivity between regions distinguish head impact exposure level in high school and youth football players. *Medical Imaging 2018: Computer-Aided Diagnosis*, *10575*, 105750F.

Nédélec, M., McCall, A., Carling, C., Legall, F., Berthoin, S., & Dupont, G. (2012). Recovery in soccer. *Sports medicine*, *42*(12), 997–1015.

Nikolaidis, P. T., Clemente, F. M., van der Linden, C. M., Rosemann, T., & Knechtle, B. (2018). Validity and reliability of 10-Hz global positioning system to assess in-line movement and change of direction. *Frontiers in physiology*, *9*, 228.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in health sciences education*, *15*(5), 625–632.

Ray, S. (2019). A quick review of machine learning algorithms. *2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon)*, 35–39.

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological methods*, *17*(3), 354.

Robitzsch, A. (2020). Why ordinal variables can (almost) always be treated as continuous variables: Clarifying assumptions of robust continuous and ordinal factor analysis estimation methods. *Frontiers in Education*, *5*, 177.

Rossi, A., Pappalardo, L., Cintia, P., Iaia, F. M., Fernández, J., & Medina, D. (2018). Effective injury forecasting in soccer with GPS training data and machine learning. *PloS one*, *13*(7), e0201264.

Rossi, A., Perri, E., Pappalardo, L., Cintia, P., & Iaia, F. M. (2019). Relationship between External and Internal Workloads in Elite Soccer Players: Comparison between Rate of Perceived Exertion and Training Load. *Applied Sciences*, *9*(23), 5174.

Sansone, P., Tschan, H., Foster, C., & Tessitore, A. (2020). Monitoring training load and perceived recovery in female basketball: Implications for training design. *The Journal of Strength & Conditioning Research*.

Saw, A. E., Main, L. C., & Gastin, P. B. (2016). Monitoring the athlete training response: Subjective self-reported measures trump commonly used objective measures: A systematic review. *British journal of sports medicine*, *50*(5), 281–291.

Sawczuk, T., Jones, B., Scantlebury, S., & Till, K. (2018). Relationships between training load, sleep duration, and daily well-being and recovery measures in youth athletes. *Pediatric exercise science*, *30*(3), 345–352.

Selmi, O., Gonçalves, B., Ouergui, I., Sampaio, J., & Bouassida, A. (2018). Influence of well-being variables and recovery state in physical enjoyment of professional soccer players during small-sided games. *Research in Sports Medicine*, *26*(2), 199–210.

Selmi, O., Ouergui, I., Castellano, J., Levitt, D., & Bouassida, A. (2020). Effect of an intensified training period on well-being indices, recovery and psychological aspects in professional soccer players. *European Review of Applied Psychology*, *70*(6), 100603.

Shan, G., Zhang, H., & Jiang, T. (2020). Correlation coefficients for a study with repeated measures. *Computational and mathematical methods in medicine*, *2020*.

Singh, D., & Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, *97*, 105524.

Tessitore, A., Meeusen, R., Cortis, C., & Capranica, L. (2007). Effects of different recovery interventions on anaerobic performances following preseason soccer training. *The Journal of Strength & Conditioning Research*, *21*(3), 745–750.

Thorpe, R. T., Strudwick, A. J., Buchheit, M., Atkinson, G., Drust, B., & Gregson, W. (2015). Monitoring fatigue during the in-season competitive phase in elite soccer players. *International journal of sports physiology and performance*, *10*(8), 958–964.

Thorpe, R. T., Strudwick, A. J., Buchheit, M., Atkinson, G., Drust, B., & Gregson, W. (2017). The influence of changes in acute training load on daily sensitivity of morning-measured fatigue variables in elite soccer players. *International journal of sports physiology and performance*, *12*(s2), S2-107-S2-113.

Vescovi, J. D., Klas, A., & Mandic, I. (2019). Investigating the relationships between load and recovery in women's field hockey–Female Athletes in Motion (FAiM) study. *International Journal of Performance Analysis in Sport*, *19*(5), 672–682.

Zhang, C.-X., Wang, G.-W., & Zhang, J.-S. (2012). An empirical bias–variance analysis of DECORATE ensemble method at different training sample sizes. *Journal of Applied Statistics*, *39*(4), 829–850.