RESEARCH QUARTERLY FOR EXERCISE AND SPORT
https://doi.org/10.1080/02701367.2024.2358956

Elite Soccer Players' Weekly Workload Assessment Through a New Training Load and Performance Score

Guglielmo Pillitteri [a], Alessio Rossi [b], Tindaro Bongiovanni [c], Giuseppe Puleo [d], Marco Petrucci [d], F Marcello Iaia [e], Hugo Sarmento [f], Filipe Manuel Clemente [g,h,i], and Giuseppe Battaglia [a]

[a]University of Palermo; [b]Feel Good Plus S.R.L. - MyPowerSet; [c]University of Bologna; [d]Palermo Football Club; [e]Università degli Studi di Milano; [f]University of Coimbra; [g]Instituto Politécnico de Viana do Castelo; [h]Sport Physical Activity and Health Research & Innovation Center; [i]Instituto de Telecomunicações, Gdansk University of Physical Education and Sport

ABSTRACT

Purpose: Monitoring players' training load allows practitioners to enhance physical performance while reducing injury risk. The aim of this study was to identify the key external load indicators in professional U19 soccer. **Methods**: Twenty-four-professional Italian young (U19) soccer players were monitored by using the rating of perceived exertion (CR-10 RPE scale) and a wearable inertial sensor during the competitive season. Three main components were detected by a Principal Component Analysis (PCA): i) volume metabolic related component, ii) intensity mechanical stimuli component, and iii) intensity metabolic/mechanical component. We hence computed two scores (i.e. Performance [PERF] and total workload [WORK]) permitting to investigate the weekly microcycle. **Results**: Correlation analysis showed that scores (i.e. PERF and WORK) are low correlated ($r = -0.20$) suggesting that they were independent. Autocorrelation analysis showed that a weekly microcycle is detectable in all the scores. Two-way ANOVA RM showed a statistical difference between match day (MD) and playing position for the three PCA components and PERF score. **Conclusion**: We proposed an innovative approach to assess both the players' physical performance and training load by using a machine learning approach allowing reducing a large dataset in an objective way. This approach may help practitioners to prescribe the training in the microcycle based on the two scores.

ARTICLE HISTORY
Received 25 August 2023
Accepted 20 May 2024

KEYWORDS
Football; performance; training load; training microcycle

Physical training, when is appropriately delivered in accordance with the training goal, induces psychophysiological responses that lead soccer players to reach specific adaptations. In order to elicit a desired athlete's training response, practitioners are used to monitor acute and chronic training effect. Actually, a conceptual framework carried out by Jeffries et al. (2021), have reported several training effect categories such as performance, physiological, subjective, and other (cognitive, biomechanical, etc.) measures highlighting the lack of a gold standard which includes surrogate (proxy) measures that reflect the training effect construct. The identification of the measurable components and their role within the training process is a paramount on- field practice (i.e., training monitoring) that allows practitioners to adjust and adapt the training periodization program. Noteworthy, training monitoring is crucial to identifying an athlete's adaptation to a training stimuli, as well as minimizing the risk of nonfunctional overreaching (Gabbett et al., 2017; Halson, 2014).

The training load (TL) in the context of athletic training has been described as "the input variable that is manipulated to elicit the desired training response" (F. Impellizzeri et al., 2019). Actually, monitoring of TL in soccer involves multiple measures derived from both external (i.e., the physical activities performed by players derived from training plan) and internal loads (i.e., the actual athlete's psychophysiological response that initiates to cope with the requirements elicited by the external load) (F. Impellizzeri et al., 2019). Actually, several indicators have been used to monitor soccer players considering both physiological (e.g., heart rate, maximal oxygen consumption and lactate blood concentration) (Rampinini et al., 2007a; Stølen et al., 2005) and match-demand-related components. Recently, this information has been derived from Electronic Performance and Tracking Systems (EPTS) (Maughan et al., 2021) devices such as Global Positioning System (GPS) (Pillitteri et al., 2023b; Varley et al., 2017) and inertial measurement unit (IMU) (Pillitteri et al., 2023a). Noteworthy, in order to better understand the TL to which soccer players are subjected an integrative approach that considers both external and internal loads is guaranteed by conceptual frameworks (F. Impellizzeri et al., 2019). Indeed, each TL indicator provides only a partial component of soccer TL (Pillitteri et al., 2023a).

Given a lack of gold standard to assess TL and considering the multifactorial nature of physical soccer performance, practitioners should consider several indicators to holistic overview of training effect (Pillitteri et al., 2023b). Actually, it is a tough challenge for the practitioners the management of a big amount of data and the appropriate interpretation of all the variables provided by the EPTS devices (Rojas-Valverde et al., 2019). Hence, may be useful and necessary the identification

CONTACT Alessio Rossi alessio.rossi2@gmail.com Department of Research and Development, Feel Good Plus S.R.L. - MyPowerSet, Via delle alghe 70, Rome 00126, Italy.

and selection of the key performance indicators (Oliva-Lozano et al., 2023) through specific statistical methods to reduce the dimensionality of datasets (Oliva-Lozano et al., 2023; Parmar et al., 2018; Rojas-Valverde et al., 2019, 2020) with the aim of synthesize the soccer physical performance in order to reduce possible communication problems within staff team (F. M. Impellizzeri et al., 2020).

Some previous studies have proposed new indexes and ratio that permits to understand the relationship between external and internal load indicators (Derbidge et al., 2022; Grünbichler et al., 2020) and evaluate the players' injury risk. Actually, as suggested by Impellizzeri et al. TL framework (2019), the uncoupling between internal and external load may be used to identify how an athlete is coping with their training program. Moreover, when players perform the same external load (e.g., total distance) with a lower internal load (e.g., average HR), would be assumed to reflect increased fitness level. On the other hand, an increase in internal response when the same amount of external load is performed may be attributed to a fatigued condition or a loss in fitness level. For example, the efficient index (Effindex) given by the ratio between mean speed in m·min^{-1} (i.e., external load) divided mean exercise intensity (%HRmax) (i.e., internal load) has been implemented to quantify the match stimuli dose-response (i.e., a physiological response for a given external activity performed) (Arrones et al., 2014; Torreño et al., 2016). Other studies have investigated the relationship between TL and injuries by using the ratio between acute and chronic workload (ACWR) (Bowen et al., 2020) suggesting the protective benefits of high chronic loads and physical strength against injury by highlighting to avoiding sudden spikes in acute TL (Griffin et al., 2020).

To the best of our knowledge, a global training/match load score has never been implemented considering both external and internal load. Although studies have used external/internal load ratio to identify fitness-fatigue conditions, machine learning methods have not been considered for this purpose. Therefore, the aims of this studies were to 1) obtain the key external load indicators in professional U19 soccer through principal component analysis (PCA) starting from several data provided by ISD; 2) calculate both a total training/match load score and an athlete's physical performance score considering both external and internal load indicators; 3) analyze the training microcycle (distance between two official matches [MD system]) through the scores calculated from the external and internal load indicators also considering the differences between playing positions.

Methods

Experimental approach to the problem

In this cross-sectional study, the internal and external training workloads of twenty-four-professional Italian young (U19) soccer players were monitored by using the rating of perceived exertion (CR-10 RPE scale) and a wearable inertial sensor, respectively, during the competitive season from August to February (2022–2023). Soccer players were categorized into six groups according to their role as follows: central back (CB), external strikers (ES), fullback (FB), midfielder (MD), wide midfielder (WM), and Strikers (S).

Subjects

Twenty-four-professional soccer players (age: 18,7 ± 0,5 years, height: 177 ± 6 cm, weight: 67,5 ± 7,7 Kg) competing in the U19 Italian Championship were analyzed. In detail, 1,280 observations, (53.3 ± 28.9 for each player) (109 training session and 12 Official Matches) have been assessed monitoring both external and internal load. Participants' playing positions were the following: Central Back (CB; $n = 6$), Full Back (FB; $n = 2$), Central Midfielder (MD; $n = 2$), Wide Midfielder (WM; $n = 6$), External Sticker (ES; $n = 5$), and Central Sticker (S; $n = 3$).

The following inclusion criteria were considered: 1) professional male soccer players; 2) no injury in the previous 6 months. Based on the exclusion criteria, only the goalkeepers were not eligible for the study. The total duration of the training sessions (number of training observation = 1,280) was 97 ± 20 min and mainly included technical-tactical team tasks along with some general physical exercise without ball (e.g., high-intensity interval training, HIIT). Strength and conditioning exercises were performed twice a week (i.e., one session in the gym and one session on the pitch). Moreover, the game duration was 83 ± 16 min (number of matches observation = 118). In total, 12 matches and 118 trainings (1,124 ± 113 minutes per week) were analyzed in this study. All participants signed an informed consent before taking part in the data collection. The Bioethics Committee of the University of Palermo (n. 68/2021) approved the study, which complies with the principles of the Declaration of Helsinki.

Procedures

Participants were monitored during training and official matches performed on a third-generation artificial pitch or natural grass during the 2022–2023 season. All players performed a typical 25-minute pre-match warm-up before OM and 30' before training (15' specific activation in the gym and 15' on field warm-up). Internal Load data was collected through RPE scale. TL for each participant was calculated according to the session-RPE method of Foster et al. (2001) 30 min after each session. The duration (minutes) and each RPE training session were recorded for players using smartphone applications (i.e., WhatsApp through a broadcast list). TL from each session for each player was calculated from the product of the session duration and the player's RPE (Foster et al., 2001). The training session duration included warm up, cool down, and rest between tasks.

External Load Data were collected through a wearable inertial sensor (TalentPlayers TPDev, firmware version 1.3). To avoid interunit errors, each participant was assigned the same ISD for each OM. Among the available devices, TalentPlayers was chosen as it provides very similar data compared to traditional GPS systems (i.e., instantaneous speed and distance, change of directions, and metabolic data) and it is already used by various Italian soccer teams (https://talentplayers.

com (accessed on 16 December 2016)). This ISD is a small wearable device integrating a six degrees of freedom MEMS inertial sensor, capable of providing real-time acceleration and rotation data along three orthogonal axes at a frequency of 100 Hz per channel. It is designed to be worn on the lower leg using an elastic band. The validity and reliability of the ISD have been previously reported (Pillitteri et al., 2021). All data were acquired by the TalentPlayer mobile app (software version 1.0.7) and uploaded to the TalentPlayers cloud. Table 1 reports the EL indicators used in this study.

Data analysis

Feature reduction

Principal component analysis (PCA) was used to reduce the feature dimension space in order to aggregate features to summarize the training workloads performed by soccer players during training and matches. PCA is a linear dimensionality reduction technique, which converts a set of correlated features in the high-dimensional space into a series of uncorrelated features in the low-dimensional space.

In particular, linear dimensionality reduction using Singular Value Decomposition (SVD) of the data was used to project the data to a lower dimensional space. To detect the best number of components the explained variance threshold was set at 90%. The input workload features were previously normalized between 0 and 1 to reduce intra-feature variability. 0.2 vas set as minimal limit for extracting feature PCA scores. Sklearn python library was used for PCA analysis. Kaiser–Meyer–Olkin (KMO) Test was used to measure the suitability of data for factor analysis. It determines the adequacy for each observed variable and for the complete model. KMO estimates the proportion of variance among the entire observed variable. Lower proportion is more suitable for factor analysis. KMO values range between 0 and 1. Value of KMO less than 0.6 is considered inadequate.

Scores computation

The sum of the product between PCA weight of each component and the workload performed by a player in a specific training permit to create n independent workload values as shown in Equation 1:

$$PCA\ score_i = \sum_{f=0}^{n} w_{f_i} * x_f \quad (1)$$

where i, f, w and x refer to the PCA component, the workload feature, the PCA weight and the workload value, respectively.

Due to the fact that, in accordance with the PCA weights and feature values, the PCA scores could be negative, the minimum possible value, i.e., when all the workload values are equal to 0, were added to all the PCA scores.

Based on the n PCA scores and the internal loads (expressed as rating of perceived exertion—RPE), 2 metrics were computed. In particular, one metric reflects the player performance (PERF—Equation 2), while the second one is focused on workload (WORK—Equation 3).

$$PERF = \frac{(PCA_i + \ldots + PCA_n)/n}{RPE} \quad (2)$$

$$WORK = \frac{PCA_i + \ldots + PCA_n}{n} * RPE * duration \quad (3)$$

Statistical analysis

The normality of data distribution was tested by Shapiro–Wilks' Normality test. Parametric or no-parametric statistical test was selected in accordance with the result of this test. Spearman Rho's correlation analysis (rs) was performed to identify relationships among scores. In order to meet any assumption for the correlation analysis, a repeated measure model with autoregressive approach (for time-series analysis) was used. An autocorrelation analysis was conducted in order to detect similarity between observations among the

Table 1. Descriptions of the External Load indicators.

Indicators	Type	Description (Unit of measure)
TD	Cinematic/volume	Total distance covered (m)
MS*	Cinematic/intensity	Maximum speed reached (even for <1 second)
N°INTACC*	Mechanical/volume	Number of intense accelerations >2 m/s^2
N°INTDEC*	Mechanical/volume	Number of intense decelerations >2 m/s^2
N°HSR*	Cinematic/volume	Number of High Intensity Running at >20 km/h
DHSR*	Cinematic/volume	Distance traveled at speed (>20 km/h) (m)
MP*	Metabolic/intensity	Metabolic Power (w · kg^{-1}) was calculated by multiplying EC (in J · kg^{-1} · m^{-1}) by running speed (v; in m · s^{-1}) at any given moment (i.e., every 0.2 s): P met = EC · v. In order to assess metabolic power, considering the energy expenditure and derived, the equation developed by di Prampero et al. (2005). established on previously studies by Minetti et al. (2002). and Osgnach et al. (2010). was adopted. (Watt = w).
DHMP*	Metabolic/volume	Distance traveled in the metabolic power zone (20–35 w) (m)
DEMP*	Metabolic/volume	Distance traveled in the metabolic power zone (35–55 w) (m)
DMMP*	Metabolic/volume	Distance traveled in the metabolic power zone (>55 w) (m)
N°INTACC/MIN	Mechanical/intensity	Number of intense accelerations >2 m/s^2 per minute
N°INTDEC/MIN	Mechanical/intensity	Number of intense decelerations >2 m/s^2 per minute
DHSR/MIN	Cinematic/intensity	Distance traveled in the speed zone (>20 km/h) (m) per minute
N° HSR	Cinematic/volume	Number of runs at speed zone (>20 km/h) (m)

Note. Legend: TD, Total Distance; MS, Maximal Speed; N°INTACC, Number of accelerations; N°INTDEC, Number of decelerations; N°HSR, Number of High-Speed Running; WT, Walking Time; THSR; DHSR, Distance High-Speed Running; MP, Metabolic Power; TLMP, DHMP, Distance High Metabolic Power; DEMP, Distance Elevated Metabolic Power; DMMP, Distance Max Metabolic Power; *Intensity indicator.

variable time series as a function of the time lag between them. Autocorrelation is the correlation of a signal with a delayed copy of itself as a function of delay. The analysis of autocorrelation is a mathematical tool for finding repeating patterns, such as the presence of a periodic signal obscured by noise or identifying the missing fundamental frequency in a signal implied by its harmonic frequencies. In statistics, the autocorrelation of a real or complex random process is the Pearson correlation (r) between values of the process at different times, as a function of the two times or of the time lag. Two-way analysis of variance for repeated measures on two factors (two-way ANOVA RM) based on a linear mixed model was used to detect difference among MD and role. The first aim of this analysis is to detect interaction among MD and role. The interaction indicate different response of role in MDs. Moreover, difference among roles and among MDs were investigated. The analysis was conducted by Python 3.8 programming language.

Results

Three main components were detected by PCA (KMO = 0.82). These components explained the 93.3% of the total variance among the multidimensional workloads space. The absolute values of the weights for each workload in each PCA component are provided in Figure 1. Based on these analyses, it is possible to define a name that describes the components as follows: i) volume metabolic related component, ii) intensity mechanical stimuli component, and iii) intensity metabolic/mechanical component.

The low Spearman Rho's correlation coefficients detected between PCA components indicate that these parameters were independent of each other and consequently provide different workload information. Similarly, also the two cumulative scores (i.e., PREF and WORK) were low correlated (rs = –0.20) indicating that performance and workload scores were independent. Moreover, PERF showed a moderate correlation with "Intensity mechanical stimuli"

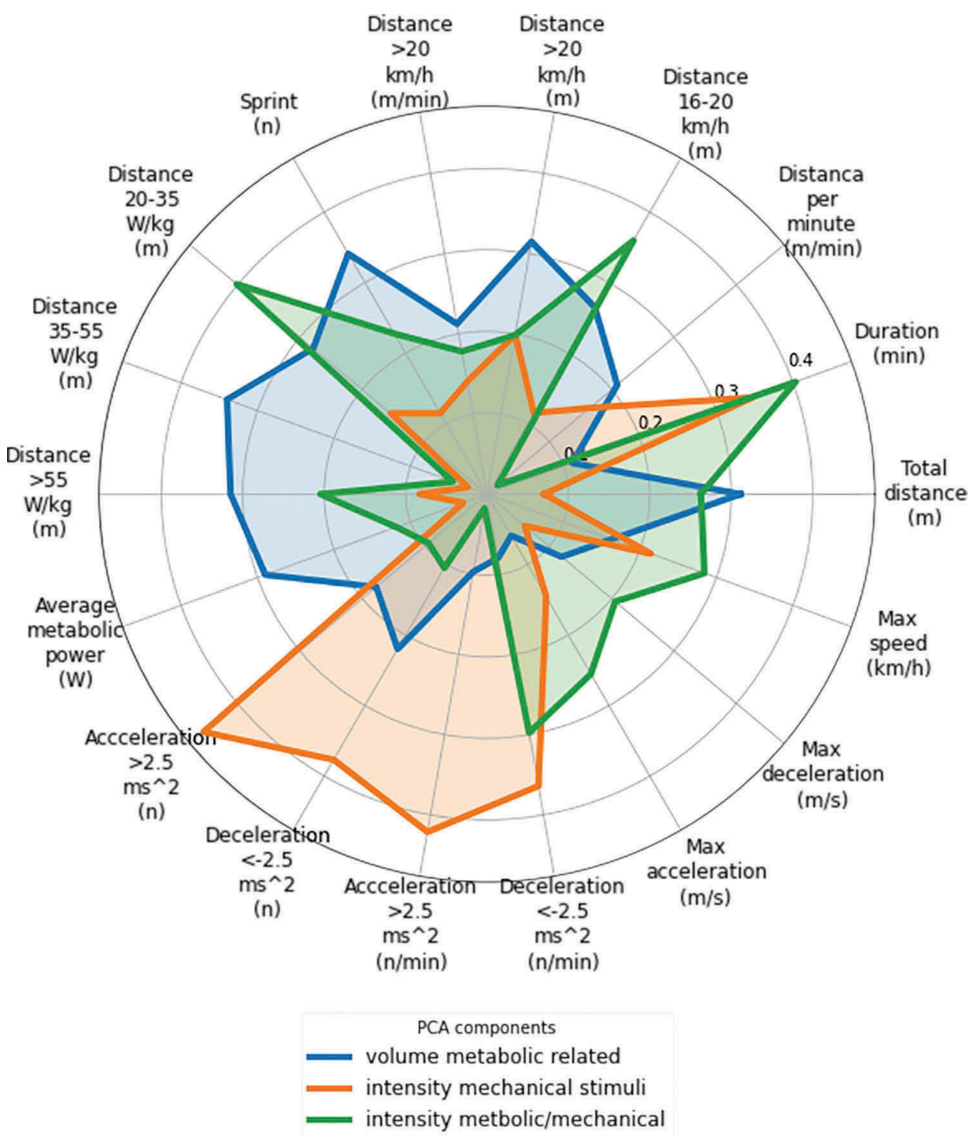

Figure 1. Absolute value for each workload feature in each PCA component. The higher the value the higher is the influence of a specific workload value on the PCA score computation.

component (rs = 0.46), while WORK with "Volume metabolic related" and "Intensity metabolic/mechanical" components (rs = 0.63 and rs = 0.34, respectively) indicating that these components were the ones that most influence the cumulative scores.

The autocorrelation analysis showed a weekly microcycle in all the scores. Actually, a positive autocorrelation was detected every 7 days and multiple of seven indicating that a similar pattern was repeated every 7 days. Moreover, a negative autocorrelation was found every 3–4 days denoting that an inverse trend was performed in the middle of the training microcycle. These results were in line with the structure of the soccer training weekly microcycle (training–recovery balance).

Two-way ANOVA RM showed a statistical difference among MD and playing position for the three PCA components and PERF (Table 2). Even if there were differences in both MD and playing position, the trend of these scores among MD were similar in each playing position. Differently, a statistical interaction was detected for WORK (Table 2) showing a different trend among playing position in different MDs.

Discussion

The main aim of this study was to create two Scores that allow to detect both individuals' physical performance and global TL (i.e., PERF and WORK scores, respectively) by using external and internal load indicators. The scores permit to analyze the microcycle workload by considering a singular indicator avoiding to control several indicators provided by MEMS devices facilitating the understanding of the workload behavior especially for the technical staff. Based on the two computed score we found a statistical difference among MDs and playing position. The scores showed a similar MD pattern for each playing position. While a statistical interaction was found for WORK score showing a different trend among playing position in different MDs.

To calculate scores, three main components, that permits to assess different characteristics of the training workloads, have been obtained by PCA as: i) volume metabolic related component, ii) intensity mechanical stimuli component, and iii) intensity metabolic/mechanical component. The combination of these components and the internal load (RPE) allows us to compute the two scores in accordance with the aim of this study. The performance score (i.e., PERF) is computed as the ratio between external and internal load (EL/IL) (Derbidge

Table 2. Descriptive statistics is reported as mean and interquartile range. The interquartile range is provided between brackets reporting the first and third interquartile values. Moreover, satistical analysis of each score splitted by role and GD.

GD	role	volume metabolic related[b]	intensity mechanical stimuli[a,b]	intensity metabolic mechanical[a,b]	PERF[b]	WORK[int]
2	CB	1.89 [1.65, 2.06]	0.45 [0.34, 0.57]	0.55 [0.26, 0.75]	0.2 [0.14, 0.23]	496.9 [366.07, 587.84]
2	CM	2.22 [2.02, 2.37]	0.42 [0.32, 0.5]	0.31 [0.2, 0.39]	0.2 [0.13, 0.19]	577.02 [404.7, 805.29]
2	ES	1.87 [1.57, 2.21]	0.4 [0.3, 0.54]	0.56 [0.32, 0.76]	0.21 [0.15, 0.25]	445.11 [335.47, 519.92]
2	FB	1.84 [1.69, 2.03]	0.45 [0.27, 0.55]	0.6 [0.4, 0.83]	0.17 [0.14, 0.18]	531.9 [432.67, 611.1]
2	S	1.87 [1.41, 2.28]	0.43 [0.27, 0.49]	0.47 [0.31, 0.65]	0.16 [0.12, 0.2]	529.01 [439.12, 625.9]
2	WM	2.02 [1.63, 2.35]	0.48 [0.31, 0.64]	0.4 [0.22, 0.46]	0.2 [0.13, 0.27]	500.69 [345.28, 668.44]
3	CB	1.85 [1.72, 2.03]	0.64 [0.42, 0.81]	0.34 [0.19, 0.47]	0.18 [0.15, 0.21]	577.87 [447.15, 734.34]
3	CM	1.94 [1.66, 2.22]	0.69 [0.62, 0.83]	0.24 [0.12, 0.34]	0.24 [0.2, 0.25]	474.66 [376.43, 601.01]
3	ES	1.91 [1.67, 2.18]	0.71 [0.46, 0.98]	0.24 [0.13, 0.31]	0.21 [0.15, 0.23]	531.83 [405.89, 618.1]
3	FB	1.73 [1.45, 2.02]	0.51 [0.29, 0.7]	0.32 [0.22, 0.38]	0.15 [0.12, 0.18]	583.55 [396.68, 735.71]
3	S	1.82 [1.71, 1.91]	0.68 [0.49, 0.83]	0.22 [0.09, 0.35]	0.17 [0.15, 0.19]	563.78 [445.95, 645.27]
3	WM	1.86 [1.61, 2.17]	0.8 [0.53, 1.09]	0.22 [0.09, 0.28]	0.27 [0.18, 0.35]	429.1 [341.36, 522.77]
4	CB	2.31 [1.93, 2.8]	0.4 [0.09, 0.62]	0.36 [0.23, 0.49]	0.19 [0.14, 0.2]	603.9 [475.49, 735.39]
4	CM	2.27 [2.06, 2.64]	0.5 [0.33, 0.6]	0.24 [0.08, 0.4]	0.2 [0.19, 0.22]	524.59 [455.51, 569.88]
4	ES	2.31 [2.1, 2.64]	0.4 [0.16, 0.58]	0.28 [0.13, 0.33]	0.2 [0.16, 0.24]	498.32 [381.39, 566.63]
4	FB	2.43 [2.1, 2.75]	0.47 [0.23, 0.68]	0.28 [0.12, 0.46]	0.16 [0.14, 0.18]	680.49 [576.93, 783.85]
4	S	2.45 [2.16, 2.91]	0.39 [0.2, 0.62]	0.2 [0.1, 0.26]	0.18 [0.17, 0.2]	550.29 [476.38, 628.33]
4	WM	2.33 [2.09, 2.73]	0.44 [0.29, 0.55]	0.24 [0.08, 0.42]	0.25 [0.19, 0.29]	447.98 [274.26, 622.62]
5	CB	1.58 [1.24, 1.86]	0.51 [0.31, 0.64]	0.32 [0.26, 0.4]	0.18 [0.14, 0.22]	399.02 [285.59, 497.03]
5	CM	2.06 [1.99, 2.28]	0.65 [0.52, 0.73]	0.12 [0.06, 0.13]	0.26 [0.22, 0.3]	379.68 [314.67, 391.52]
5	ES	1.94 [1.39, 2.51]	0.48 [0.41, 0.65]	0.19 [0.08, 0.26]	0.23 [0.15, 0.3]	354.02 [217.44, 436.17]
5	FB	1.67 [1.37, 1.88]	0.3 [0.25, 0.41]	0.32 [0.25, 0.39]	0.17 [0.12, 0.18]	391.23 [267.87, 507.56]
5	S	1.58 [1.43, 1.88]	0.39 [0.33, 0.43]	0.17 [0.08, 0.25]	0.2 [0.16, 0.22]	289.18 [205.45, 330.5]
5	WM	1.63 [1.5, 1.84]	0.44 [0.27, 0.59]	0.25 [0.16, 0.35]	0.3 [0.22, 0.41]	239.11 [172.58, 308.03]
−1	CB	0.41 [0.22, 0.5]	0.42 [0.19, 0.58]	0.12 [0.07, 0.14]	0.17 [0.13, 0.2]	38.67 [19.85, 46.92]
−1	CM	0.44 [0.4, 0.48]	0.74 [0.48, 0.87]	0.09 [0.05, 0.11]	0.32 [0.22, 0.37]	37.32 [26.55, 48.92]
−1	ES	0.59 [0.37, 0.69]	0.4 [0.27, 0.44]	0.22 [0.14, 0.26]	0.24 [0.15, 0.26]	45.04 [31.66, 57.26]
−1	FB	0.41 [0.28, 0.53]	0.28 [0.15, 0.33]	0.11 [0.04, 0.13]	0.23 [0.11, 0.27]	20.7 [12.52, 25.61]
−1	S	0.36 [0.22, 0.54]	0.41 [0.15, 0.68]	0.15 [0.05, 0.25]	0.24 [0.09, 0.31]	26.15 [12.54, 39.42]
−1	WM	0.21 [0.11, 0.3]	0.31 [0.16, 0.46]	0.06 [0.03, 0.08]	0.15 [0.06, 0.23]	16.61 [13.6, 18.01]
0	CB	1.97 [1.77, 2.29]	0.18 [0.07, 0.25]	0.34 [0.19, 0.48]	0.1 [0.09, 0.11]	647.08 [499.49, 901.32]
0	CM	1.61 [1.35, 1.69]	0.12 [0.05, 0.17]	0.19 [0.04, 0.26]	0.12 [0.11, 0.13]	282.11 [115.05, 348.49]
0	ES	2.09 [1.62, 2.48]	0.13 [0.06, 0.2]	0.14 [0.06, 0.19]	0.13 [0.12, 0.15]	419.07 [195.31, 568.62]
0	FB	2 [1.88, 2.22]	0.13 [0.06, 0.19]	0.24 [0.11, 0.35]	0.11 [0.11, 0.13]	516.86 [434.42, 699.47]
0	S	2.29 [2.16, 2.6]	0.14 [0.09, 0.19]	0.11 [0.04, 0.15]	0.1 [0.09, 0.11]	625.62 [551.77, 728]
0	WM	2.06 [1.84, 2.33]	0.1 [0.04, 0.13]	0.22 [0.14, 0.23]	0.18 [0.14, 0.2]	349.43 [242.69, 419.54]

Note. In particular, the two-way ANOVA statistical difference is reported as: [a]difference among role; [b]difference among GD; [int]interaction between role and GD.

et al., 2022). The use of this ratio to understand physical performance (i.e., athlete's fitness status) is supported by the works carried out by Gabbett et al. (2017) and Impellizzeri et al. framework (2019). Accordingly, some studies have implemented the EL/IL ratio to understand how an athlete responds to the administered load. For example, Derbidge et al. proposed three different ratios to assess the players' loads (i.e., total distance/HR exertion [HRe], player load/HRe, and equivalent distance/HRe) (Derbidge et al., 2022). Unlike previous studies, we selected the EL indicators not arbitrarily, but by a data-driven approach that permits creating more holistic and not human-biased external load indicators.

Concerning the attempt to calculate a global TL indicator, the WORK score has been computed considering EL, IL (i.e., based on RPE), and duration. The rationale behind the implementation of this score can be understood considering the TL framework (including both EL and IL) and can be operationalized following the exposure/dose—response function theory (i.e., a concept common in epidemiology that describes the relationship between the amounts of a factor/agent [i.e., dose] and the responses [e.g., risk of an outcome] (Checkoway et al., 2004; F. M. Impellizzeri et al., 2023; Kriebel et al., 2007)). In sports context, the theoretical dose—response function is explained by the parabola (or "inverted U" shape) where the performance improves with increases in TL up until a maximum level is reached, and outside this point, further increases in TL consequence in a worsening of performance (overreaching and overtraining). Actually, WORK score is calculated in accordance with TL theory as the product of Duration (i.e., time) and Intensity (i.e., calculated as EL indicators reduced by PCA and RPE for internal load). Several studies have monitored the training/match load in soccer considering EL, IL, or both EL and IL (integrated approach) as suggested by Impellizzeri framework (F. Impellizzeri et al., 2019; F. M. Impellizzeri et al., 2023). Noteworthy, no study has ever implemented a formula to obtain a global load that includes both EL (by selecting load indicators with complex analysis such as PCA) and IL. Moreover, many studies have used indicators derived from GPS or Inertial devices considering them as EL. Actually, in order to obtain the "external load," the values detected by a given indicator (e.g., Total distance) has to be multiplied with match/training duration, as suggested by the TL framework carried out by F. Impellizzeri et al. (2019). However, since none of these external load indicators has been multiplied by the duration of a given exercise or of the entire training session/competition, it does not seem possible to be able to refer to the "TL concept" (i.e., EL). In fact, Coyne et al. (2022), have recommended using the training impulse (the product of training volume and intensity factors) for IL and EL calculation, rather than a singular volume or intensity factor. This approach allows comparing "apples with apples" avoiding methodological differences among studies which could make comparisons difficult inducing possible misleading interpretation (Coyne et al., 2022). The use of a data driven approach, i.e., PCA, is paramount to reduce the large amount of available indicators as more than one indicator should be considered to obtain a complete and appropriate view of the physical performance of the soccer players taking into consideration all the components of the EL (i.e., metabolic, speed-related and accelerations-related). Indeed, authors (F. Impellizzeri et al., 2019; F. M. Impellizzeri et al., 2023; Pillitteri et al., 2023b) highlighted that a lack of a "gold standard" to determine intensity in soccer makes explaining the intensity concept difficult and, consequently, making it difficult to select appropriate indicators to assess the soccer physical performance. In this way, practitioners should consider together different EL and IL indicators during the monitoring process because each of them explains a partial component of soccer physical performance. PCA selection allows the detection of EL indicators more appropriately than a subjective approach. Actually, our analysis shows a low correlation between PCA components that are independent of each other and consequently provide different information. Consequently, the two cumulative Scores (i.e., PREF and WORK) are low correlated ($r = -0.20$) indicating that performance and workload scores provide different information despite common measures of exposure (e.g., acute and chronic TL and their ratio) lack conceptual and compu-tational validity (F. Impellizzeri et al., 2019).

In this study, PERF and WORK scores were used to evaluate the typical soccer physical performance and the microcycle workload considering also the different playing position. The timeseries analysis of the scores showed a weekly microcycle structure (i.e., a positive autocorrelation was detected every 7 days and multiple of 7 indicating that a similar pattern is repeated every 7 days). These results are in line with the structure of the soccer training weekly microcycle (training–recovery balance) (Kenttä & Hassmén, 1998; Nédélec et al., 2012; Selmi et al., 2022). For example, Oliva-Lozano et al. (2023), found a significant difference in load variability within the microcycle depending on the training day. Actually, previous studies have also showed that the TL decreases as the training day is closest (e.g., greater workload in −4MD or −3MD compared to −2MD or −1MD) (Clemente et al., 2019; Martín-García et al., 2018). However, to the best of the authors' knowledge, only Oliva-Lozano et al. (2023), have analyzed the load variability within the microcycle considering EL key indicators through PCA showing the lowest variability of all training days observed in −1MD (due to tapering strategies adopted by coaches) while the highest in +1MD (Fessi et al., 2016). These results are corroborated by the current study that found the highest EL (based on scores) in MD + 3 and MD + 4 and the lowest one in MD-1.

One of the novel findings of our study was the investigation of the microcycle structure (i.e., MD system) load based on the new scores computed considering also playing position. Literature shows a different load for each playing position during training and competition (DiSalvo et al., 2006, 2013; Pillitteri, Giustino, et al., 2023c; Rampinini et al., 2007b). In the same way, we have found a statistical difference among MD and among playing positions for the three PCA components and PERF score, but they follow the same training microcycle structure. Conversely, Nobari et al. (2023), reported that the weekly TL changed throughout the season, although no significant differences were observed between playing positions. These results disagree with our finding and despite the trend of these scores across MDs was similar in each playing position, a statistical interaction was detected for WORK score showing a different trend among playing positions in different MDs (i.e., each playing position follows a well-defined WORK's microcycle). The study of TL is complex, and the conclusion should be contextualized to the sample examined. Several variables can contribute to determine the external and internal load such as, training status, methodology, level of competition and training typology. Indeed, studies should deeply provide specific information about training (e.g., training type, recovery strategies, athlete fitness status) to better understand the athlete's physiological and perception responses to training (i.e., internal load) as shown by previous studies on RPE (Lupo et al., 2016).

Conclusion

This study aimed to create two scores from external and internal load indicators to assess both individuals' physical performance (i.e., PERF score) and global training workload (i.e., WORK score). PCA has provided three different characteristics of the training external workloads (i.e., i) volume metabolic related component, ii) intensity mechanical stimuli component, and iii) intensity metabolic/mechanical component). The combination of these components and the internal load (RPE) allows us to compute the two scores. The Performance score (i.e., PERF) is computed as the ratio between external and internal load (EL/IL), while the Work score has been computed considering EL, IL (i.e., based on RPE), and duration. Moreover, a weekly microcycle structure has been provided through the timeseries analysis showing that a pattern is repeated every 7 days for both scores. Also, by analyzing the scores across MDs, a similar result in each playing position have been found, although the WORK score shows a different trend among playing positions in different MDs (i.e., each playing position present a well-defined WORK's microcycle).

This paper would like to propose an innovative approach to assess both the players' physical performance and TL by using a machine learning approach that permits to reduce a large dataset considering the main external load indicators in an objective way. The main strength of this work is the use of the Performance and Workload scores to monitor training and soccer matches by including the main external load indicators, the players' physiological responses (i.e., internal load) and duration. Our approach allows to assess the overall training/match load and players' performance considering only two indicators against numerous indicators that could be redundant and selected on the basis of the subjective choices of the practitioners. This study presents some limitations. Data has been collected in young male professional soccer players with a number of both training and matches observed within August and February. Accordingly, care should be taken when generalizing these findings due to the relatively limited number of observations in the total sample. Despite several external load indicators have been included in our analysis, the variables related to speed and acceleration as well as metabolic power were set on absolute thresholds (e.g., distance >20 Km/h for high-speed running), and other metric with different thresholds could have been included. In this regard, future studies may consider the addition of individualized zones based on maximum speed or maximum acceleration capacities. Several external (environmental) and internal factors (psychological, stress, sleep, fatigue) could affect the physical performance during matches and training. For example, the tactical requirements of coaches (e.g., 1-4-3-3 vs −1-3-5-2 system) determine different external load responses. Future studies should include a larger sample, including other levels of competition, tactical requirement, as well as other sports to strength the power of PCA analysis. Although the Score aiming to synthetized both physical performance and training load useful for coaches, given the fundamental contribution (with unique and specific meaning and information) provided by each external and internal load indicators, the monitoring process should include both Scores and both external and internal load indicators for more in-depth analysis.

Practical application

The machine learning approach conducted in this study provides valuable information on training/match load and players physical performance and relative playing position in the microcycle structure (i.e., MD system) based on the two computed scores. As a matter of fact, in order to describe the players' physical performance and workload during training and matches two scores have been computed from EL and IL data. Actually, the approach used in this work allows practitioners to better identify the workloads performed by players during training microcycle or official match along with player's performance. Since practitioners should consider together different EL and IL indicators during the monitoring process because each of them explains a partial component of soccer physical performance, PCA analysis allows the detection of EL indicators more appropriately than a subjective approach. In detail, this approach permits to detect relevant information from load and performance. Therefore, it is possible to understand the differences between training and official matches or between MDs, considering also each player's playing position. In this way, practitioners could prescribe the training drill in the microcycle considering the results provided by two scores.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

No funding was received for this paper

ORCID

Guglielmo Pillitteri http://orcid.org/0000-0002-0905-6401
Alessio Rossi http://orcid.org/0000-0002-6400-5914

References

Arrones, L. S., Torreno, N., Requena, B., De Villarreal, E., Casamichana, D., Carlos, J., & Barbero-Alvarez, D. (2014). Match-play activity profile in professional soccer players during official games and the relationship between external and internal load. *Journal of Sports Medicine and Physical Fitness*, *55*, 1417–1422. https://doi.org/10.1123/ijspp.2015-0252

Bowen, L., Gross, A. S., Gimpel, M., Bruce-Low, S., & Li, F.-X. (2020). Spikes in acute: Chronic workload ratio (ACWR) associated with a 5–7 times greater injury rate in English Premier League football players: A comprehensive 3-year study. *British Journal of Sports Medicine*, *54*(12), 731–738. https://doi.org/10.1136/bjsports-2018-099422

Checkoway, H., Pearce, N., & Kriebel, D. (2004). *Research methods in occupational epidemiology (Vol. 34)*. Monographs in Epidemiology and.

Clemente, F. M., Owen, A., Serra-Olivares, J., Nikolaidis, P. T., Van Der Linden, C. M., & Mendes, B. (2019). Characterization of the weekly external load profile of professional soccer teams from Portugal and the Netherlands. *Journal of Human Kinetics*, *66*(1), 155–164. https://doi.org/10.2478/hukin-2018-0054

Coyne, J. O., Coutts, A. J., Newton, R. U., & Haff, G. G. (2022). The current state of subjective training load monitoring: Follow-up and future directions. *Sports Medicine - Open*, *8*(1), 53. https://doi.org/10.1186/s40798-022-00433-y

Derbidge, M. C., Moody, J., & Byrne, P. (2022). Can the use of an external: Internal ratio during standardised small-sided games be a detector of fatigue in professional adolescent soccer players?.

diPrampero, P. E., Fusi, S., Sepulcri, L., Morin, J. B., Belli, A., & Antonutto, G. (2005). Sprint running: A new energetic approach. *The Journal of Experimental Biology*, *208*(Pt 14), 2809–2816. https://doi.org/10.1242/jeb.01700

DiSalvo, V., Baron, R., Tschan, H., Montero, F. C., Bachl, N., & Pigozzi, F. (2006). Performance characteristics according to playing position in elite soccer. *International Journal of Sports Medicine*, *28*(3), 222–227. https://doi.org/10.1055/s-2006-924294

DiSalvo, V., Pigozzi, F., Gonzalez-Haro, C., Laughlin, M., & De Witt, J. (2013). Match performance comparison in top English soccer leagues. *International Journal of Sports Medicine*, *34*(6), 526–532. https://doi.org/10.1055/s-0032-1327660

Fessi, M. S., Zarrouk, N., DiSalvo, V., Filetti, C., Barker, A. R., & Moalla, W. (2016). Effects of tapering on physical match activities in professional soccer players. *Journal of Sports Sciences*, *34*(24), 2189–2194. https://doi.org/10.1080/02640414.2016.1171891

Foster, C., Florhaug, J. A., Franklin, J., Gottschall, L., Hrovatin, L. A., Parker, S., Doleshal, P., & Dodge, C. (2001). A new approach to monitoring exercise training. *The Journal of Strength & Conditioning Research*, *15*(1), 109–115. https://doi.org/10.1519/00124278-200102000-00019

Gabbett, T. J., Nassis, G. P., Oetter, E., Pretorius, J., Johnston, N., Medina, D., Rodas, G., Myslinski, T., Howells, D., Beard, A., & Ryan, A. (2017). The athlete monitoring cycle: A practical guide to interpreting and applying training monitoring data. *British Journal of Sports Medicine*, *51*(20), 1451–1452. BMJ Publishing Group Ltd and British Association of Sport and Exercise Medicine. https://doi.org/10.1136/bjsports-2016-097298

Griffin, A., Kenny, I. C., Comyns, T. M., & Lyons, M. J. (2020). The association between the acute: Chronic workload ratio and injury and its application in team sports: A systematic review. *Sports Medicine*, *50*(3), 561–580. https://doi.org/10.1007/s40279-019-01218-2

Grünbichler, J., Federolf, P., & Gatterer, H. J. (2020). Workload efficiency as a new tool to describe external and internal competitive match load of a professional soccer team: A descriptive study on the relationship between pre-game training loads and relative match load. *European Journal of Sport Science*, *20*(8), 1034–1041. https://doi.org/10.1080/17461391.2019.1697374

Halson, S. L. J. S. M. (2014). Monitoring training load to understand fatigue in athletes. *Sports Medicine*, *44*(Suppl 2), 139–147. https://doi.org/10.1007/s40279-014-0253-z

Impellizzeri, F. M., Marcora, S. M., & Coutts, A. J. (2019). Internal and external training load: 15 years on. *International Journal of Sports Physiology and Performance*, *14*(2), 270–273. https://doi.org/10.1123/ijspp.2018-0935

Impellizzeri, F. M., McCall, A., Ward, P., Bornn, L., & Coutts, A. (2020). Training load and its role in injury prevention, part 2: Conceptual and methodologic pitfalls. *Journal of Athletic Training*, *55*(9), 893–901. https://doi.org/10.4085/1062-6050-501-19

Impellizzeri, F. M., Shrier, I., McLaren, S. J., Coutts, A. J., McCall, A., Slattery, K., Jeffries, A. C., & Kalkhoven, J. T. (2023). Understanding training load as exposure and dose. *Sports Medicine*, 1–13. https://doi.org/10.1007/s40279-023-01833-0

Impellizzeri, F., Wookcock, S., McCall, A., Ward, P., & Coutts, A. J. (2019). *The acute-chronic workload ratio-injury figure and its 'sweet spot'are flawed.*

Jeffries, A. C., Marcora, S. M., Coutts, A. J., Wallace, L., McCall, A., & Impellizzeri, F. M. (2021). Development of a revised conceptual framework of physical training for use in research and practice. *Sports Medicine*, 1–16. https://doi.org/10.1007/s40279-021-01551-5

Kenttä, G., & Hassmén, P. J. S. M. (1998). Overtraining and recovery: A conceptual model. *Sports Medicine*, *26*(1), 1–16. https://doi.org/10.2165/00007256-199826010-00001

Kriebel, D., Checkoway, H., & Pearce, N. (2007). Exposure and dose modelling in occupational epidemiology. *Occupational and Environmental Medicine*, *64*(7), 492–498. https://doi.org/10.1136/oem.2006.030031

Lupo, C., Capranica, L., Cortis, C., Guidotti, F., Bianco, A., & Tessitore, A. (2016). Session-RPE for quantifying load of different youth taekwondo training sessions. *The Journal of Sports Medicine and Physical Fitness*, *57*(3), 189–194. https://doi.org/10.23736/S0022-4707.16.06021-X

Martín-García, A., Díaz, A. G., Bradley, P. S., Morera, F., & Casamichana, D. (2018). Quantification of a professional football team's external load using a microcycle structure. *The Journal of Strength & Conditioning Research*, *32*(12), 3511–3518. https://doi.org/10.1519/JSC.0000000000002816

Maughan, P., Swinton, P., & MacFarlane, N. (2021). Relationships between training load variables in professional youth football players. *International Journal of Sports Medicine*, *42*(7), 624–629. https://doi.org/10.1055/a-1300-2959

Minetti, A. E., Moia, C., Roi, G. S., Susta, D., & Ferretti, G. (2002). Energy cost of walking and running at extreme uphill and downhill slopes. *Journal of Applied Physiology*, *93*(3), 1039–1046. https://doi.org/10.1152/japplphysiol.01177.2001

Nédélec, M., McCall, A., Carling, C., Legall, F., Berthoin, S., & Dupont, G. (2012). Recovery in soccer: Part I - post-match fatigue and time course of recovery. *Sports Medicine*, *42*(12), 997–1015. https://doi.org/10.1007/BF03262308

Nobari, H., Praça, G. M., da Glória Teles Bredt, S., González, P. P., Clemente, F. M., Carlos-Vivas, J., & Ardigò, L. P. (2023). Weekly variations of accelerometer variables and workload of professional soccer players from different positions throughout a season. *Scientific Reports*, *13*(1), 2625. https://doi.org/10.1038/s41598-023-29793-5

Oliva-Lozano, J. M., Barbier, X., Fortes, V., & Muyor, J. M. (2023). Key load indicators and load variability in professional soccer players: A full season study. *Research in Sports Medicine*, *31*(3), 201–213. https://doi.org/10.1080/15438627.2021.1954517

Osgnach, C., Poser, S., Bernardini, R., Rinaldo, R., & diPrampero, P. E. (2010). Energy cost and metabolic power in elite soccer: A new match analysis approach. *Medicine and Science in Sports and Exercise*, *42*(1), 170–178. https://doi.org/10.1249/MSS.0b013e3181ae5cfd

Parmar, N., James, N., Hearne, G., & Jones, B. J. (2018). Using principal component analysis to develop performance indicators in professional rugby league. *International Journal of Performance Analysis in Sport*, *18*(6), 938–949. https://doi.org/10.1080/24748668.2018.1528525

Pillitteri, G., Clemente, F. M., Petrucci, M., Rossi, A., Bellafiore, M., Bianco, A., Palma, A., & Battaglia, G. (2023a). Toward a New Conceptual Approach to "Intensity" in Soccer Player's Monitoring: A Narrative Review. *The Journal of Strength & Conditioning Research* *37*(9), 1896–1911. https://doi.org/10.1519/JSC.0000000000004503

Pillitteri, G., Giustino, V., Petrucci, M., Rossi, A., Leale, I., Bellafiore, M., Thomas, E., Iovane, A., Palma, A., & Battaglia, G. (2023c). Match load physical demands in U-19 professional soccer players assessed by a wearable inertial sensor. *Journal of Functional Morphology and Kinesiology*, *8*(1), 22. https://doi.org/10.3390/jfmk8010022

Pillitteri, G., Rossi, A., Simonelli, C., Leale, I., Giustino, V., & Battaglia, G. (2023b). Association between internal load responses and recovery ability in U19 professional soccer players: A machine learning approach. *Heliyon*, *9*(4), e15454. https://doi.org/10.1016/j.heliyon.2023.e15454

Pillitteri, G., Thomas, E., Battaglia, G., Navarra, G. A., Scardina, A., Gammino, V., Ricchiari, D., & Bellafiore, M. J. S. (2021). Validity and reliability of an inertial sensor device for specific running patterns in soccer. *Sensors*, *21*(21), 7255. https://doi.org/10.3390/s21217255

Rampinini, E., Coutts, A. J., Castagna, C., Sassi, R., & Impellizzeri, F. (2007a). Variation in top level soccer match performance. *International Journal of Sports Medicine*, *28*(12), 1018–1024. https://doi.org/10.1055/s-2007-965158

Rampinini, E., Impellizzeri, F. M., Castagna, C., Abt, G., Chamari, K., Sassi, A., & Marcora, S. M. (2007b). Factors influencing physiological responses to small-sided soccer games. *Journal of Sports Sciences*, *25*(6), 659–666. https://doi.org/10.1080/02640410600811858

Rojas-Valverde, D., Gómez-Carmona, C. D., Gutiérrez-Vargas, R., & Pino-Ortega, J. (2019). From big data mining to technical sport reports: The case of inertial measurement units. *BMJ Open Sport and Exercise Medicine*, *5*(1), e000565. https://doi.org/10.1136/bmjsem-2019-000565

Rojas-Valverde, D., Pino-Ortega, J., Gómez-Carmona, C. D., & Rico-González, M. (2020). A systematic review of methods and criteria standard proposal for the use of principal component analysis in team's sports science. *International Journal of Environmental Research Public Health*, *17*(23), 8712. https://doi.org/10.3390/ijerph17238712

Selmi, O., Ouergui, I., Muscella, A., My, G., Marsigliante, S., Nobari, H., Suzuki, K., & Bouassida, A. (2022). Monitoring psychometric states of recovery to improve performance in soccer players: A brief review. *International Journal of Environmental Research and Public Health*, *19*(15), 9385. https://doi.org/10.3390/ijerph19159385

Stølen, T., Chamari, K., Castagna, C., & Wisløff, U. (2005). Physiology of soccer: An update. *Sports Medicine*, *35*(6), 501–536. https://doi.org/10.2165/00007256-200535060-00004

Torreño, N., Munguía-Izquierdo, D., Coutts, A., de Villarreal, E. S., Asian-Clemente, J., & Suarez-Arrones, L. (2016). Relationship between external and internal loads of professional soccer players during full matches in official games using global positioning systems and heart-rate technology. *International Journal of Sports Physiology and Performance*, *11*(7), 940–946. https://doi.org/10.1123/ijspp.2015-0252

Varley, M. C., Jaspers, A., Helsen, W. F., & Malone, J. J. (2017). Methodological considerations when quantifying high-intensity efforts in team sport using global positioning system technology. *International Journal of Sports Physiology & Performance*, *12*(8), 1059–1068. https://doi.org/10.1123/ijspp.2016-0534