MAUGHAN, P.C., MACFARLANE, N.G. and SWINTON, P.A. [2021]. The influence of season phase on multivariate load relationships in professional youth soccer. *Journal of sports sciences* [online], (accepted). To be made available from: https://doi.org/10.1080/02640414.2021.1993642

The influence of season phase on multivariate load relationships in professional youth soccer.

MAUGHAN, P.C., MACFARLANE, N.G. and SWINTON, P.A.

2021

The Influence of Season Phase on Multivariate Load Relationships in Professional Youth Soccer

Journal of Sports Sciences

Authors: Patrick C Maughan [a,b,c], Niall G MacFarlane [b], Paul A Swinton [c]

[a] Aberdeen Football Club

Cormack Park

Kingswells

Aberdeen

AB15 8SP

[b] University of Glasgow

College of Medical and Veterinary Life Sciences

Human Biology/Sports Science

West Medical Building

Glasgow

G12 8QQ

[c] Robert Gordon University

School of Health Sciences

Ishbel Gordon Building

Aberdeen

AB10 7QE

Abstract

The purpose of this research was to assess the relationships between subjective and external measures of training load in professional youth footballers, whilst accounting for the effect of the stage of the season. Data for ratings of perceived exertion (RPE) and seven global positioning systems (GPS) derived measures were collected from 20 players (age = 17.4 ± 1.3 yrs, height = 178.0 ± 8.1 cm, mass = 71.8 ± 7.2 kg) across a 47-week season. The season was categorised by a pre-season phase, and two competitive phases (Comp1, Comp2). The structure of the data were investigated using principal component analysis. An extraction criterion of component with eigenvalues ≥ 1.0 was used. Two components were retained for the pre-season period explaining a cumulative variance of 77.1%. Single components were retained for both Comp1 and Comp2 explaining 73.3% and 74.3% of variance, respectively. Identification of single components may suggest that measures are related and can be used interchangeably, however these interpretations should be considered with caution. The identification of multiple components in the pre-season phase suggests that univariate measures may not be sufficient when considering load experienced. These results suggest that factoring load based on measures of volume and intensity should be considered.

Keywords: Training load, monitoring, session-rpe, team sports, workload

Introduction

Soccer match play is characterised by frequent high intensity accelerations, decelerations, and running [1]. As such, soccer training aims to prepare players for the physical demands of match play, alongside developing technical, tactical and psychological understanding. Due to the high physical demands involved, match play and training to prepare soccer players can also present substantive risk of injury [2]. With the aim of improving performance, and reducing the risk of injury, practitioners supporting professional soccer players routinely monitor the physical load experienced by players [3]. Whilst this route of investigation is common, it has been suggested that current practices relating load monitoring with injury are lacking in substantial evidence, possibly due to the shortcomings of available univariate load metrics [4] Load and the subsequent adaptations generated, can be characterised as being either physiological or biomechanical [5]. Features of training load describing the magnitude and amount of the physical work are considered the external load [5,6], whereas, features describing the resultant physiological and biomechanical response are characterised as the internal load [5,6]. Generally, practitioners monitor prescribed physical work, which is represented by external load, alongside the players response which is characterised as the internal load [5,6]. A central aim of research is to accurately model relationships between external and internal load to create more effective and responsive training stimuli to enhance physical performance and its expression during match play [7].

A range of technologies, variables, data processing and analysis techniques are used when monitoring internal and external load. Common approaches to monitor internal load include subjective measurements such as the rating of perceived exertion (RPE) and objective measurements including heart-rate (HR) based assessments in the form of training impulse (TRIMP) and time spent in specific HR zones [8]. Development of technologies such as global position system (GPS) devices and accelerometers has increased the availability of external load variables which are now common in professional soccer [9]. Whilst advances in technology and greater dissemination of research-based practices has made continuous load monitoring an essential component of elite athlete support, the lack of criterion measures of load has led practitioners to collect a range of variables posing a challenge to clear interpretation of the data [10]. Initial attempts to assess validity of outcomes or identify underlying structures to reduce the dimensionality of data have been achieved by comparing all measures against each other using correlation or principal component approaches, respectively [10]. Research investigating underlying structure has generally found that measures representing either the internal or external load are strongly related to each other [10]. However, research has also established that relationships between load monitoring variables may be influenced by different training modes [10-13]. Comparing research findings across different sports suggests that potential changes in underlying structure across different training modes may also be sport specific [11,12]. Previous research in rugby league showed significant effects of training mode on relationships between internal and external load measures [10]. Similar findings were found in a follow up study in rugby league comparing relationships between load measures during skills and conditioning focused training sessions [11]. In contrast, a recent analysis in professional youth soccer found no changes in underlying structure when categorising training sessions based on their proximity to match day (e.g., MD-1, MD-2) [12]. In accordance with previous research, the structure of load measures aligned themselves along measures of volume and intensity [13]. It is plausible that the contrasting results may be influenced by the specificity of the training sessions, where mode of training is more clearly defined in rugby league and sessions can be categorised for example as ‘skills’ or ‘conditioning’ [11]. Conversely in soccer training, there is often less specificity and sessions are generally categorised based on their proximity to match day creating greater within-session variability and potentially masking more subtle changes in relationships [12].

Whilst preliminary evidence suggest that load relationships remain consistent across different training contexts in professional soccer, less is known about the effect of stage of season. Previous research investigating training load in professional soccer has compared internal and external load in the English Premier League [14]. Malone et al. [14] reported no significant differences across the pre-season and in-season phases of training; however, it is worth noting that match play data was not included which may have the potential to influence overall load experienced, particularly during the in-season phase [14]. The aims of the different phases of the season are generally different, with development of fitness a primary goal of pre-season [14] and often maintenance of previously developed physical qualities the aim during in-season to enable focus on technical and tactical development [14]. Given the contrasting aims of different stages of the season, there is potential that the underlying structure described by the multivariate relationships between load measures may also change. As it is routine for practitioners to collect many load variables without criterion, greater understanding of underlying structure and the factors that can alter this will provide practitioners with better context to monitor players throughout the season. Therefore, the aim of the current study was to quantify and describe the relationship between internal and external load variables across phases of the season. Specifically, we aimed to assess the relationship between sRPE and various external load measures collected via GPS technology. To do this the study used analyses methods previously used to assess the underlying structure of relationships [10-13].

Methods

Subjects

Data were collected from 20 male professional youth soccer players (age 17.4 ± 1.3 yrs, height 178.0 ± 8.1 cm, mass 71.8 ± 7.2 kg). All data were collected during the 2018/19 season. Data comprised players from multiple positions, but data provided from goalkeepers were removed. In accordance with previous research [14], data recorded from a small selection of non-representative training sessions were removed to limit the influence of outliers. Post-Match top-ups, rehabilitation sessions, and non-pitch-based sessions such as resistance training were also excluded from the analysis. As the aim of this study was to compare different phases of the season, the winter break period was not included in the analyses.

Design

The present study employed a prospective design with data collection across a 47-week season with Scottish professional youth soccer players. The data collection periods comprised a 6-week pre-season and two competitive phases lasting 20 weeks (Comp1) and 19 weeks (Comp2), respectively. The competitive phases were split by a 2-week winter break. Subjective measures of training load were collected via RPE. Objective measures of training load were collected via commercially available GPS units. Data were collected for all training sessions and matches. Data collected and the retrospective nature of the data analysis conformed to the University of Glasgow research policies and were in accordance with the Declaration of Helsinki.

Methodology

RPE was collected, in isolation, approximately 30 minutes after each training session using a commonly utilised modified BORG-CR10 scale [8,15] that had been used extensively with players previous to the study. Each RPE score was multiplied by session duration to obtain subjective training load [16]. Alongside this measurement of subjective training load, objective external training load was also collected. Players wore commercially available GPS units (Optimeye X4, Firmware version 7.27; Catapult Sports, Melbourne, Australia) previously used in research conducted in team sports [11,17]. The units include a GPS receiver and a triaxial accelerometer collecting data at 10 Hz and 100 Hz, respectively. Velocity and acceleration dwell times were set at 0.6 and 0.4 s, respectively. As per previous recommendations, each player wore the same device for each session [18]. Following training or matches, data were downloaded and analysed via the Openfield software package (Software version 1.19, Catapult Sports). Average satellite count was 10.6 ± 1.7. The average horizontal dilution of precision (HDOP) was 0.8 ± 0.2. Variables selected to quantify external load were total distance (m), PlayerLoad (au), low intensity running (<14.4km.h^{-1}, m) high-speed running distance (19.8 – 24.98 km.h^{-1}, m) sprinting distance (>24.98km.h^{-1}, m), accelerations (>2m.s^{-2}, count) and decelerations (> -2m.s^{-2}, count).

Statistical Analysis

Following previously described procedures [12] we carried out a correlation analysis before performing principal component analysis (PCA) on each stage of season. Where data were missing, they were treated as missing at random and imputed using the MICE package in the R statistical environment (version 4.0.3; R Foundation for Statistical Computing, Vienna, Austria.) [19]. Relationships between all load variables were quantified during each stage of season using Pearson's product moment correlation. Following this, data were prepared for PCA by firstly visually inspecting the correlation matrix to assess the factorability of the dataset [20]. The suitability of data were then assessed using the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy, and the Bartlett test of sphericity [21]. KMO (~chi square) values were 0.76 (5187.241), 0.84 (16931.8), and 0.83 (16078.5) for Pre-Season, Comp1 and Comp2, respectively. All tests of sphericity were significant ($p<0.001$). A KMO value of 0.5 or above has previously been identified as a suitable result to perform PCA [22,23] and has been used in similar research [11,12]. PCA was carried out using the 'prcomp' function of the R stats package (v3.6.2) [24] and the 'principal' function of the psych package (v2.0.12) [25]. Principal components with an eigenvalue ≥1.0 were retained for extraction [23]. When two or more principal components were retained based on their eigenvalue, varimax rotation was performed. For each retained principal component, only the original load variables with a principal component loading of >0.7 were retained [22].

Results

There were 3207 individual recordings included in the analysis comprising 695 individual MD recording and 2512 individual training session recording. Distribution of the mean loads during each phase of the season are presented in Table 1. Correlations including 95% confidence intervals for each phase of season are presented in Figure 1. Total distance, PlayerLoad and low-intensity running showed very-large correlations ($r \geq 0.77$) across all phases of the season. High-speed running distance showed moderate to very-large correlations ($0.39 \leq r \leq 0.70$), whilst sprinting distance showed moderate correlations across the season ($0.32 \leq r \leq 0.45$). Finally, accelerations showed large correlations across all phases ($r \geq 0.52$), whilst decelerations showed large to very-large correlations ($0.54 \leq r \geq 0.75$).

Results of the PCA are presented in Tables 2 and 3. Two principal components were identified for pre-season whilst one component was identified for each competitive phase. Variance explained and loadings are presented for the pre-season phase following varimax rotation. The components explained 77.1% of the variance for the pre-season phase. The un-rotated principal components for Comp1 and Comp 2 explained 73.3% and 74.3% of the variance, respectively. The heaviest component loadings for Comp1 and Comp2 were total distance (Comp1 = 0.96, Comp2 = 0.95), PlayerLoad (Comp1 = 0.94, Comp2 = 0.95) and low intensity running (Comp1 = 0.93, Comp2 = 0.93).

*****Insert Tables 1,2 and 3 about here*****

*****Insert Figure 1 about here*****

Discussion

The primary finding of this study was the identification of multiple components during the pre-season period, and conversely the identification of a single component within both competitive phases. This finding suggests in the pre-season phase univariate assessments of load may be insufficient when characterising the load experienced by players. [11,12] Conversely, the identification of a single component with relatively similar loadings across all variables obtained during both competitive phases suggest that load measures may be used interchangeably.

Previous research in professional rugby league [10,11] and in professional soccer [12] has reported that multiple measures are required to capture the variance across different training themes when expressed as training mode, or relative to match day. In each of these studies, two or more components were identified following PCA. To our knowledge this is first assessment of this relationship when considering the phase of the season. In the present study the pre-season stage produced two components and following varimax rotation, the component loadings could be described as representative of either training volume or intensity [12]. In the present study, PCA carried out on pre-season data produced two principal components which represented 77.11% of the cumulative variance. The highest rotated component loadings for component one were sRPE (0.85), total distance (0.9), PlayerLoad (0.91) and low-intensity running (0.94). For rotated component two, the highest loadings were high-speed running (0.79), sprinting (0.87) and acceleration (0.57). Studies in rugby league have shown that variables generally align based on categories of internal or external training load [10,11]. In the present study we only included sRPE as a measure of subjective internal load. This may have influenced our findings, however, there does still seem to be some relationship between measures which may provide similar information regarding either volume or intensity of training or match play.

Whilst our analysis produced multiple principal components when investigating the pre-season phase, we only identified one component when analysing both competitive phases. This would suggest that all load variables fit into one theoretical factor, and could, theoretically, be used interchangeably [10]. It is worth noting that this may be due to the method we selected for defining how many components would be retained for rotation. A recent review concerning the use of PCA in sport found that 62.2% of the studies analysed retained factors for rotation if they had an eigenvalue >1 [26]. Other methods, such as visual analysis of an eigenvalue scree plot whereby the 'elbow' of the data would be identified [20], may have led to retention of two principal components for competitive phase data. Had we included a second factor in both analyses then the results would have been comparable to our presented pre-season data (Table 2). Retention of two factors for Comp1 would have resulted in two principal components which would have explained 84.6% of the variance. Rotated component loadings would also have corresponded with our pre-season findings. Factor loadings for the first rotated component would have been 0.88, 0.9, 0.88 and 0.94 for sRPE, total distance, PlayerLoad and low-intensity running, respectively. The second rotated component would again have been best represented by high-speed running (0.77), sprinting (0.93), accelerations (0.63) and additionally decelerations (0.61). Similarly, for Comp2, retention of two factors would have results in a cumulative variance explained of 84.4%. Rotated component loadings would also have been similar to pre-season findings. Component 1 would have been best represented by sRPE (0.88), total distance (0.91), PlayerLoad (0.92), and low-intensity running (0.94). Component 2 would again have been best represented by high-speed running (0.68) and sprinting (0.94). Interestingly loadings for accelerations and decelerations were slightly lower than may have been presented for Comp1 with values of 0.47 and 0.58 respectively. Clearly the method selected by practitioners for retaining factors will effect results, with the most popular method used currently in practice being the Kaiser criteria (eigenvalue >1) [20].

The findings from the present study alongside previous work [12] demonstrates that sRPE is representative of a measure of volume. Previous research has shown that both RPE and sRPE are significantly related to several external load and intensity measures [27,28]. When analysing youth soccer players, the strongest within-individual correlations between sRPE and various external load measures were found for duration (r = 0.767), distance (r = 0.699) and distance in acceleration (r = 0.696) [28]. Using generalized estimating equation (GEE) models, it was found that PlayerLoad, high-speed distance and distance in acceleration were the strongest contributory variables when estimating sRPE [28]. However, in our present study it is worth noting the strong component loadings of acceleration and deceleration within the first rotated component of each analyses, which may suggest that subjective perception of effort, may also be strongly related to measures of acceleration and deceleration, but not high-speed running or sprinting.

The findings of the present study further evidence that measures of sRPE appears to provide information regarding load volume, rather than intensity. Practitioners should consider this when analysing this measure to represent the load experienced by athletes. Whilst our analysis shows that this relationship is not consistent across stages of the season, this is likely due to retention criteria applied. Therefore, practitioners should consider the stage of the season, and the physical goals of that phase, when assessing load measurements.

The findings of the present study should be interpreted given the following limitations of the research. The categorisation method used in the present study comprised three levels for analysis and a logical comparison between a pre-season phase, and two competitive phases. However, future analysis may wish to investigate shorter mesocycle periods within the competitive period, for example 6-week blocks, to provide a more in-depth comparison across the season. Additionally, the present study did not attempt to differentiate structure of load variables across different categories of players of players. Further differentiation in terms of partitioning within and between variance in structure, or potential differences across for example starters, non-starters, or fringe players, may also provide additional insight to the proposed relationships. Additionally, the present study only included one subjective measure of internal load due to player adherence with objective methods, such as heart-rate based measures. Further insight to objective measures of internal load may provide useful insight regarding previously observed relationships between internal and external measures of load [10].

This study provides further evidence that univariate measures may not be sufficient when measuring the load experienced by players and that this limitation may be influenced by factors such as the stage of the season. These results, alongside previous results, would suggest that factoring load based on measures of volume and intensity would be appropriate. Whilst analyses of both competitive phases of the season identified only one principal component, which would suggest that variables may be used interchangeably during this period, it is worth noting that the criteria selected for retaining factors plays a key role in this process. As previously suggested, the dose-response relationship with changes in fitness, or injury occurrence, for these combined load measures should be a future aim of analyses.

References

1. Whitehead S, Till K, Weaving D, Jones B. The use of microtechnology to quantify the peak match demands of the football codes: a systematic review. *Sports Med.* 2018;48(11):2549-2575.

2. Peterson L, Junge A, Chomiak J, Graf-Baumann T, Dvorak J. Incidence of football injuries and complaints in different age groups and skill-level groups. *AM J Sports Med.* 2000;28(5_suppl):51-57.
3. Drew MK, Finch CF. The relationship between training load and injury, illness and soreness: a systematic and literature review. *Sports Med.* 2016;46(6):861-883.
4. Kalkhoven JT, Watsford ML, Coutts AJ, Edwards WB, Impellizzeri FM. Training load and injury: causal pathways and future directions. *Sports Med.* 2021:1-14.
5. Vanrenterghem J, Nedergaard NJ, Robinson MA, Drust B. Training load monitoring in team sports: a novel framework separating physiological and biomechanical load-adaptation pathways. *Sports Med.* 2017;47(11):2135-2142.
6. Impellizzeri FM, Marcora SM, Coutts AJ. Internal and external training load: 15 years on. *Int J Sports Physiol Perform.* 2019;14(2):270-273.
7. Halson SL. Monitoring training load to understand fatigue in athletes. *Sports Med.* 2014;44(2):139-147.
8. Impellizzeri FM, Rampinini E, Coutts AJ, Sassi A, Marcora SM. Use of RPE-based training load in soccer. *Med Sci Sports Exerc.* 2004;36(6):1042-1047.
9. Akenhead R, Nassis GP. Training load and player monitoring in high-level football: current practice and perceptions. *Int J Sports Physiol Perform.* 2016;11(5):587-593.
10. Weaving D, Marshall P, Earle K, Nevill A, Abt G. Combining internal-and external-training-load measures in professional rugby league. *Int J Sports Physiol Perform.* 2014;9(6):905-912.
11. Weaving D, Jones B, Marshall P, Till K, Abt G. Multiple measures are needed to quantify training loads in professional rugby league. *Int J Sports Med.* 2017;38(10):735-740.
12. Maughan P, Swinton P, MacFarlane N. Relationships between training load variables in professional youth football players. *Int J Sports Med.* 2021;42(7):624-629.
13. Maughan PC, MacFarlane NG, Swinton PA. Relationship Between Subjective and External Training Load Variables in Youth Soccer Players. *Int J Sports Physiol Perform.* 2021;1(aop):1-7.
14. Malone JJ, Di Michele R, Morgans R, Burgess D, Morton JP, Drust B. Seasonal training-load quantification in elite English premier league soccer players. *Int J Sports Physiol Perform.* 2015;10(4):489-497.
15. Foster C, Hector LL, Welsh R, Schrager M, Green MA, Snyder AC. Effects of specific versus cross-training on running performance. *Eur J Appl Physiol Occup Physiol.* 1995;70(4):367-372.
16. Foster C, Florhaug JA, Franklin J, et al. A new approach to monitoring exercise training. *J Strength Cond Res.* 2001;15(1):109-115.

17. Jones RN, Greig M, Mawéné Y, Barrow J, Page RM. The influence of short-term fixture congestion on position specific match running performance and external loading patterns in English professional soccer. *J Sports Sci.* 2019;37(12):1338-1346.
18. Scott MT, Scott TJ, Kelly VG. The validity and reliability of global positioning systems in team sport: a brief review. *J Strength Cond Res.* 2016;30(5):1470-1490.
19. Buuren Sv, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *J Stat Softw.* 2010:1-68.
20. Tabachnick BG, Fidell LS, Ullman JB. *Using multivariate statistics.* Vol 5: Pearson Boston, MA; 2007.
21. Bartlett MS. A note on the multiplying factors for various χ2 approximations. *J R Stat Soc.* 1954;16(2):296-298.
22. Hair JF, Black WC, Babin BJ, Anderson RE, Tatham R. Multivariate data analysis . Uppersaddle River. NJ: Pearson Prentice Hall; 2006.
23. Kaiser HF. The application of electronic computers to factor analysis. *Educ Psychol Meas.* 1960;20(1):141-151.
24. Team RC. R: A language and environment for statistical computing. 2013.
25. Revelle W, Revelle MW. Package 'psych'. *The comprehensive R archive network.* 2015;337:338.
26. Rojas-Valverde D, Pino-Ortega J, Gómez-Carmona CD, Rico-González M. A Systematic Review of Methods and Criteria Standard Proposal for the Use of Principal Component Analysis in Team's Sports Science. *Int J Environ Res Public Health.* 2020;17(23):8712.
27. Gaudino P, Iaia FM, Strudwick AJ, et al. Factors influencing perception of effort (session rating of perceived exertion) during elite soccer training. *Int J Sports Physiol Perform.* 2015;10(7):860-864.
28. Marynowicz J, Kikut K, Lango M, Horna D, Andrzejewski M. Relationship Between the Session-RPE and External Measures of Training Load in Youth Soccer Training. *J Strength Cond Res.* 2020;34(10):2800-2804.

	Duration (mins)	sRPE (au)	Total Distance (m)	PlayerLoad (au)	LIR (m)	HSR (m)	Sprinting (m)	Accel (count)	Decel (count)
Pre-Season	57.8 ± 17.8	360 ± 191	4861 ± 2175	525 ± 220	3929 ± 1610	213 ± 246	37.1 ± 58.3	20.7 ± 12.2	14.5 ± 9.57
Comp1	64 ± 19.7	369 ± 200	5361 ± 2444	594 ± 251	4495 ± 1857	186 ± 181	46.9 ± 73.7	23.0 ± 11.5	16.5 ± 9.98
Comp2	60.3 ± 21.3	357 ± 215	5263 ± 2717	565 ± 275	4356 ± 2055	194 ± 185	48 ± 65	22.2 ± 11.4	16.3 ± 10.3

Table 1 – Mean (± SD) duration and load measures across phase of season. LIR, Low intensity running; HSR, High speed running; Accel, Accelerations; Decel, Decelerations

	Pre-Season	
	Principal Component	
	1	2
Eigenvalue	5.11	1.06
% of Variance	63.9	13.21
Cumulative Variance %	63.9	77.11
	Rotated Component	
	1	2
% of Variance	51.14	77.11
	Rotated Component Loadings	
	1	2
sRPE	**0.85**	0.18
Total Distance	**0.9**	0.32
PlayerLoad	**0.91**	0.31
LI.Running	**0.94**	0.18
Running	0.26	**0.79**
Sprinting	0.16	**0.87**
Accelerations	0.53	0.57
Decelerations	0.69	0.33

Table 2 – PCA results for Pre-Season phase. This includes the eigenvalue, and % of variance explained. LIR, low intensity running; HSR, high speed running; Accel, accelerations; Decel, decelerations

	Comp1	Comp2
	Principal Component	
	1	1
Eigenvalue	5.86	5.95
% of Variance	73.25	74.32
	Component Loadings	
	1	1
sRPE	**0.86**	**0.91**
Total Distance	**0.96**	**0.95**
PlayerLoad	**0.94**	**0.95**
LIR	**0.93**	**0.93**
HSR	**0.84**	**0.85**
Sprinting	0.67	0.64
Accel	**0.74**	**0.73**
Decel	**0.88**	**0.88**

Table 3 – PCA results for Comp1 & Comp2. This includes the eigenvalue, and % of variance explained. As these phases did not meet retention criteria for further components, only the un-rotated values for the first principal component are presented. LIR, low intensity running; HSR, high speed running; Accel, accelerations; Decel, decelerations

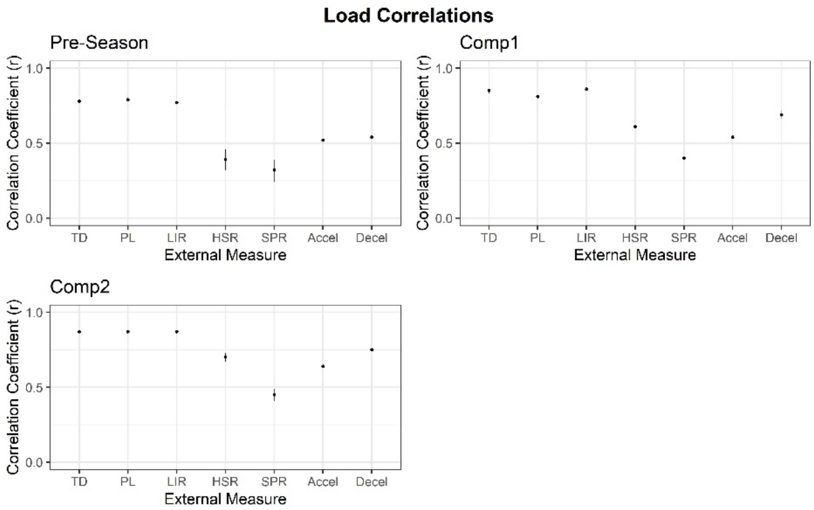

Figure 1 - Pearson's product moment correlations between sRPE and all external load measures (error bars represent 95% CI). TD, Total Distance; PL, PlayerLoad; LIR, low intensity running; HSR, running; SPR, sprinting; Accel, accelerations; Decel, decelerations.