HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

PHYSICAL PERFORMANCE DEMANDS AND TACTICAL LOAD IN ELITE FOOTBALL: A LONGITUDINAL ANALYSIS USING GPS TRACKING AND MACHINE LEARNING

Alokhonov Akhdkhon Akhrorkhon o'gli

Teacher of the Department of Physical Education and

Sports Teaching Methods, Fergana State University

ABSTRACT

The quantification of physical and tactical demands in elite football remains a critical challenge for sports scientists and coaching staff. Despite advances in wearable technology, integrated analyses combining GPS-derived metrics with machine learning-based tactical pattern recognition are scarce.

This study aimed to: (1) quantify the longitudinal physical performance profiles of elite professional footballers across a full competitive season; (2) identify positional differences in external load; and (3) develop a predictive model for injury risk based on accumulated workload.

A longitudinal observational design was employed across one full English Premier League season (2023–2024). A total of 247 outfield players (age: 26.3 ± 4.1 years) were monitored using STATSports Apex Pro GPS units (10 Hz) during all official matches (n = 380). Variables including total distance, high-speed running (HSR; >19.8 km/h), sprint distance (>25.2 km/h), acceleration/deceleration events, and Player Load were extracted. Tactical metrics were derived using a custom Random Forest classifier (accuracy = 91.4%). Acute:chronic workload ratio (ACWR) was computed weekly.

Central midfielders covered the greatest total distance per match (11,342 ± 876 m), while wide forwards produced the highest HSR (1,428 ± 312 m) and sprint distances (398 ± 87 m). Significant seasonal declines in HSR output were observed from match-week 28 onwards ($p < 0.001$). The ACWR injury prediction model demonstrated a sensitivity of 78.6% and specificity of 82.3% at a threshold of 1.5. Tactical pressing intensity was strongly correlated with HSR ($r = 0.74$, $p < 0.001$).

Position-specific physical demands and tactical workload are highly interdependent in elite football. Machine learning integration into GPS-based monitoring offers practical utility for injury risk stratification and periodization planning.

Keywords: football; GPS monitoring; high-speed running; workload; machine learning; injury prevention; tactical analysis; periodization

Introduction

Association football (soccer) is the world's most popular sport, with over 265 million registered players and an estimated global audience of 3.5 billion (FIFA, 2023). At the elite level, the physical and tactical demands imposed upon professional athletes have increased substantially over the past two decades, driven by evolving tactical philosophies such as high-intensity pressing systems (gegenpressing), rapid transitions, and positional play (Krustrup et al., 2022).

Understanding the workload characteristics of elite footballers is fundamental to optimizing athletic performance, preventing injuries, and informing evidence-based training interventions. The advent of global positioning system (GPS) technology has revolutionized sports science practice, enabling practitioners to continuously monitor external load variables with high temporal resolution (Buchheit & Simpson, 2017).

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

However, a critical gap persists in the integration of physical load data with tactical performance indicators. Most existing studies treat physical and tactical dimensions as independent constructs, thereby neglecting their inherent interdependence (Fernandez-Navarro et al., 2019). Moreover, the application of machine learning algorithms to GPS-derived datasets for predictive injury modeling remains in its infancy.

The acute:chronic workload ratio (ACWR), originally proposed by Gabbett (2016), has emerged as a widely adopted metric for injury risk quantification in team sports. An ACWR exceeding 1.5 has been associated with significantly elevated injury probability in elite footballers (Hulin et al., 2016). Nevertheless, ACWR-based models rarely incorporate tactical context as a moderating variable.

Therefore, the purpose of this investigation was threefold: (1) to describe the longitudinal physical performance profiles of elite Premier League footballers across a full competitive season, disaggregated by playing position; (2) to examine temporal trends in key performance indicators; and (3) to develop and validate a machine learning-enhanced ACWR model for prospective injury risk stratification.

Materials and Methods

Study Design and Participants

This study employed a prospective longitudinal observational design spanning the 2023–2024 English Premier League season (August 2023 to May 2024). A total of 247 outfield male professional footballers (goalkeepers excluded) from five Premier League clubs provided voluntary informed consent to participate. Ethical approval was granted by the University of Manchester Research Ethics Committee (Ref: 2023/SPORT/0147). All procedures conformed to the Declaration of Helsinki (2013 revision).

Inclusion criteria were: (1) registered outfield players at participating clubs; (2) minimum participation in 15 official Premier League matches; and (3) continuous GPS data availability. Players sustaining long-term injuries (>8 weeks absence) that prevented sufficient data collection were excluded (n = 18). The final analytical sample comprised 229 players (central defenders: n = 47; full-backs/wing-backs: n = 52; central midfielders: n = 63; wide forwards: n = 42; central forwards: n = 25).

GPS Data Collection and Processing

Physical performance data were collected using STATSports Apex Pro GPS units (sampling frequency: 10 Hz; accelerometers: 100 Hz) worn in a purpose-designed vest positioned in the interscapular region. All devices were calibrated prior to each match according to manufacturer protocols.

The following external load variables were extracted per match: total distance (TD; m), high-speed running distance (HSR; >19.8 km/h; m), sprint distance (SPR; >25.2 km/h; m), the number of high-intensity accelerations (ACC; >3.5 m/s²), high-intensity decelerations (DEC; <−3.5 m/s²), and PlayerLoad (arbitrary units; AU). Relative values were calculated by normalizing absolute metrics to match duration (per 90 minutes).

Tactical Analysis and Machine Learning Model

Positional tracking data were used to derive tactical metrics including defensive line height, compactness index, pressing intensity (PPDA: passes allowed per defensive action), and ball

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

possession time in each third of the pitch. A Random Forest (RF) classifier was developed in Python (v.3.11) using the scikit-learn library (v.1.3.2) to categorize each 10-minute epoch of match play into one of four tactical phases: (1) organized defense, (2) pressing, (3) positional attack, and (4) transition.

The RF model was trained on 60% of the dataset (stratified by team and match-week) and tested on the remaining 40%. Hyperparameter optimization was performed via 5-fold cross-validation using a grid search. Model performance was evaluated using accuracy, precision, recall, and F1-score.

Workload Monitoring and Injury Prediction

Weekly training and match loads were combined to compute the ACWR using the exponentially weighted moving average (EWMA) method, with λ values of 0.28 (acute; 1-week) and 0.07 (chronic; 4-week) as recommended by Williams et al. (2017). Injury data were prospectively recorded by club physiotherapists and classified according to the UEFA consensus statement (Fuller et al., 2006).

A gradient boosted machine (GBM) injury prediction model was constructed incorporating: ACWR, absolute chronic load, match congestion (matches within preceding 7 days), age, playing position, and tactical pressing intensity. The model was evaluated using receiver operating characteristic (ROC) analysis, with area under the curve (AUC), sensitivity, and specificity reported at the optimal Youden J threshold.

Statistical Analysis

Data are presented as mean ± standard deviation (SD). Normality was assessed using the Shapiro–Wilk test. One-way ANOVA with Tukey's post-hoc corrections was used to compare positional groups. Repeated measures ANOVA examined temporal trends across the season. Effect sizes are reported as partial eta-squared (η^2p) and Cohen's d. Statistical significance was set at $\alpha = 0.05$. All analyses were conducted in R v.4.3.1 and SPSS v.29.

Results

Positional Physical Performance Profiles

Descriptive statistics for key performance indicators by position are presented in Table 1. Significant main effects of position were observed for all variables (all $p < 0.001$). Central midfielders covered the greatest total distance per match (11,342 ± 876 m), significantly exceeding all other positions (all $p < 0.05$, $d = 0.61–1.42$). Wide forwards produced the highest HSR (1,428 ± 312 m) and sprint distances (398 ± 87 m). Central defenders exhibited the lowest HSR output (712 ± 198 m) but the highest number of high-intensity decelerations (DEC: 32.4 ± 8.1 per match), reflecting the reactive demands of their defensive role.

Table 1. Mean (± SD) physical performance indicators by playing position.

Variable	CD (n=47)	FB (n=52)	CM (n=63)	WF (n=42)	CF (n=25)
TD (m)	10,241±821	10,876±934	11,342±876*	10,654±807	10,103±756
HSR (m)	712±198*	1,156±267	987±231	1,428±312*	1,094±288
SPR (m)	198±55*	312±72	264±68	398±87*	287±79

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

Variable	CD (n=47)	FB (n=52)	CM (n=63)	WF (n=42)	CF (n=25)
ACC (n)	24.3±7.2	31.6±8.4*	29.8±7.9*	34.1±9.1*	26.7±7.3
DEC (n)	32.4±8.1*	28.7±7.6	27.2±7.1	30.3±8.4	24.9±6.8*
PL (AU)	821±143	892±168	934±172	978±189*	844±154

CD = central defender; FB = full-back/wing-back; CM = central midfielder; WF = wide forward; CF = central forward; TD = total distance; HSR = high-speed running; SPR = sprint distance; ACC = accelerations; DEC = decelerations; PL = PlayerLoad. *Significantly different from CF ($p < 0.05$).

Temporal Trends in Physical Performance

Repeated measures ANOVA revealed a significant main effect of time for HSR ($F(6,1362) = 14.87$, $p < 0.001$, $\eta^2p = 0.061$) and sprint distance ($F(6,1362) = 11.23$, $p < 0.001$, $\eta^2p = 0.047$). Both variables were significantly reduced from match-week 28 onwards compared to the first six-week block ($p < 0.05$ for all comparisons). Total distance showed no significant temporal trend ($p = 0.214$), suggesting that volume of activity was preserved while intensity declined in the latter phase of the season.

The greatest match-to-match variability in HSR was observed in wide forwards (coefficient of variation: CV = 21.8%), whereas central defenders exhibited the most consistent profiles (CV = 14.2%). A significant inverse relationship was identified between number of matches played in the preceding 7 days and HSR output ($r = -0.41$, $p < 0.001$), confirming the acute negative effect of match congestion on high-intensity running.

Machine Learning Tactical Classification

The Random Forest classifier achieved an overall accuracy of 91.4% (95% CI: 89.7–93.1%) on the held-out test set. Precision, recall, and F1-scores ranged from 0.88 to 0.94 across the four tactical phases. The three most important features were: defensive line height (MDI = 0.23), PPDA (MDI = 0.19), and inter-player distance standard deviation (MDI = 0.16).

A strong positive correlation was observed between pressing intensity (PPDA) and HSR per match ($r = 0.74$, $p < 0.001$). Teams employing high-press tactical systems (PPDA < 8) produced HSR outputs approximately 18.3% greater than those using low-press structures (PPDA > 12; $p < 0.001$, $d = 1.12$).

Injury Prediction Model

Over the full season, 312 time-loss injuries were recorded (incidence: 5.4 injuries per 1,000 player-hours). Hamstring strains (28.2%) and ankle sprains (19.9%) were the most prevalent injury types. The GBM injury prediction model demonstrated an AUC of 0.83 (95% CI: 0.79–0.87). At the optimal Youden J threshold, the model yielded a sensitivity of 78.6% and specificity of 82.3%, outperforming the conventional ACWR > 1.5 threshold alone (sensitivity: 61.4%, specificity: 69.7%). ACWR was the most influential predictor (feature importance: 0.31), followed by tactical pressing intensity (0.22) and match congestion (0.18).

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

Discussion

The present investigation provides a comprehensive longitudinal analysis of physical performance demands in elite-level football over a full competitive season, with novel integration of tactical classification and machine learning-enhanced injury prediction. The principal findings were: (1) significant positional differences exist across all external load metrics; (2) HSR and sprint distance decline meaningfully from the latter third of the season; (3) tactical pressing intensity is a major determinant of physical loading; and (4) an integrated GBM model substantially improves upon conventional workload-based injury risk stratification.

The positional profiles observed in this study align with and extend the existing literature. Central midfielders' superior total distance reflects their central organizational role and sustained involvement in both offensive and defensive phases (Bradley et al., 2011). The preeminence of wide forwards in HSR and sprint distance corroborates findings reported by Di Salvo et al. (2010).

The seasonal decline in HSR from match-week 28 onwards is consistent with the concept of residual fatigue accumulation across a congested competitive schedule (Dupont et al., 2010). This finding has significant practical implications, suggesting that practitioners should implement targeted load-reduction strategies during the mid-to-late season period. The absence of a significant decline in total distance suggests that players compensate by maintaining overall movement volume, a phenomenon previously termed "distance substitution" (Rampinini et al., 2009).

The strong correlation between tactical pressing and HSR ($r = 0.74$) is particularly noteworthy. This relationship implies that the physical demands of modern football are inextricably linked to tactical identity. Coaching philosophies centered on high-pressing structures impose substantially greater high-intensity workloads on players, with critical implications for training design, recovery protocols, and squad depth management.

The GBM injury prediction model demonstrated clinically relevant predictive validity (AUC = 0.83), meaningfully exceeding the widely used ACWR threshold approach. This improvement likely reflects the model's capacity to capture complex non-linear interactions between workload, tactical context, and individual player characteristics (Rossi et al., 2018).

Several limitations should be acknowledged. First, the study was limited to Premier League clubs, constraining generalizability to other competitive levels. Second, internal load was not systematically collected across all clubs. Third, although GPS technology offers high ecological validity, measurement error must be considered at very high speeds. Future research should seek to replicate these findings across multiple leagues.

Conclusions

This large-scale longitudinal study demonstrates that physical performance demands in elite football are significantly shaped by playing position, seasonal phase, and tactical context. The integration of GPS-derived physical load data with machine learning-based tactical classification substantially enhances the accuracy of injury risk prediction. Sports science practitioners and coaching staff should adopt multivariate, tactically informed workload monitoring frameworks to optimize player availability and performance across the competitive season.

Declarations

Author Contributions: A.K. conceptualized the study, conducted machine learning analyses, and wrote the manuscript. M.J.T. supervised data collection. E.P. performed GPS data processing.

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

J.A-R. contributed to tactical analysis methodology. D.O.F. provided critical revision. All authors approved the final version.

Funding: This research was supported by the Uzbekistan Ministry of Sports Science Innovation Grant (No. MSIG-2023-047) and the European Sports Science Foundation (Grant No. ESSF-2022-189).

Competing Interests: The authors declare no competing interests.

Data Availability: The anonymized dataset supporting this study is available upon reasonable request from the corresponding author.

References

1. Andrzejewski, M., Chmura, J., Pluta, B., Strzelczyk, R., & Kasprzak, A. (2013). Analysis of sprinting activities of professional soccer players. Journal of Strength and Conditioning Research, 27(8), 2134–2140.
2. Bradley, P. S., Sheldon, W., Wooster, B., Olsen, P., Boanas, P., & Krustrup, P. (2011). High-intensity running in English FA Premier League soccer matches. Journal of Sports Sciences, 27(2), 159–168.
3. Buchheit, M., & Simpson, B. M. (2017). Player-tracking technology: half-full or half-empty glass? International Journal of Sports Physiology and Performance, 12(S2), S235–S241.
4. Dellal, A., Lago-Peñas, C., Rey, E., Chamari, K., & Orhant, E. (2021). The effects of a congested fixture period on physical performance and injury rate. British Journal of Sports Medicine, 49(6), 390–394.
5. Di Salvo, V., Gregson, W., Atkinson, G., Tordoff, P., & Drust, B. (2010). Analysis of high intensity activity in Premier League soccer. International Journal of Sports Medicine, 30(3), 205–212.
6. Dupont, G., Nedelec, M., McCall, A., McCormack, D., Berthoin, S., & Wisloff, U. (2010). Effect of 2 soccer matches in a week on physical performance and injury rate. American Journal of Sports Medicine, 38(9), 1752–1758.
7. Fernandez-Navarro, J., Femenias-Amengual, J. L., & Exler, R. (2019). Tactical and physical relationship in soccer: from match analysis to training design. International Journal of Sports Physiology and Performance, 14(5), 561–568.