UiT The Arctic University of Norway

Faculty of Health Sciences

Activity Profiles and Training Loads of Highly Trained Female Football Players

An in-depth analysis with implications for practice

Andreas Kjæreng Winther

A dissertation for the degree of Philosophiae Doctor February 2024

Table of Contents

Acknowledgement

I would like to express my gratitude to UiT The Arctic University of Norway and Tromsø Forskningsstiftelse for providing me with the opportunity to pursue a PhD. The support and resources offered by the university have been invaluable to my academic journey.

My heartfelt thanks go out to my colleagues for their camaraderie. The positive work environment they have cultivated through dedication and teamwork has greatly enriched my experience. I will fondly remember our engaging lunch conversations.

Special appreciation goes to my main supervisor, Svein Arne, for his guidance, encouragement, and support throughout this endeavor. His mentorship has been instrumental in shaping my research and academic growth. Without you, neither my own PhD journey, nor those of the other associated with the project, would have been possible.

I am immensely thankful to all the individuals, particularly Dag, Sigurd, Ivan, and Morten, who have contributed to the writing and development of the articles associated with my research. Their collaborative efforts and insights have enriched the quality of this thesis.

To my parents and sister, I extend my deepest thanks for their unwavering love and support. Their encouragement has been a constant source of motivation for me.

Last but certainly not least, I want to express my deepest appreciation to my girlfriend, Harriet, for her unwavering support. Her presence in my life has been a source of strength and encouragement. I promise to make up for lost time by helping around the apartment more (at least I'll try!)

Abstract

Background: With the surge in popularity and competitiveness of women's football, there is a call for further research to enhance the development and health of these athletes. For practitioners, understanding the performance characteristics of these athletes could be beneficial in informing training strategies. By establishing normative data for a given performance level, practitioners can readily assess their performance relative to these benchmarks, potentially shaping training goals and targets. Additionally, with the widespread use of external load monitoring, having reference values to identify deviations from typical performances becomes increasingly useful. Finally, there is a notable gap in the literature concerning the external training load of highly trained female football players in typical training cycles and in multi-team cohorts. Considering these issues, the objective of this thesis was to describe the activity profiles and training loads of highly trained female football players.

Methods: We collected tracking data from four teams in the Norwegian premier division over the course of two seasons. In paper I, we assessed the activity profiles of highly trained female players according to playing position, while in paper II, we established reference values for unusual changes in these profiles. Paper III described the external training load in typical cycles and compared differences in external training load between starters and non-starters.

Results: Players covered mean match total distances ranging from 8934 to 10131 meters, high-speed running distances ranging from 1054 to 1894 meters, sprint distances ranging from 227 to 530 meters, acceleration distances ranging from 433 to 578 meters, deceleration distances ranging from 305 to 493 meters, and achieved peak speeds of 27 to 29 $km \cdot h^{-1}$. The largest differences in activity were observed between center-backs and the wide positions (full-backs and wide midfielders), where the latter covered greater high-speed- and- sprint distances. We also observed only trivial to small (Cohen's D_z: 0.07-0.20) decreases in activity between post-peak periods and corresponding mean match periods. The observed match-to-match variability in high-speed running distance, sprint distance, and acceleration- and deceleration distance ranged from 12 to 36%, while peak speed and total distance ranged from 4.5 to 5%. In longer cycles, the majority of load was concentrated around mid-week in a pyramid-like fashion, with minimal differentiation in the load of starters and non-starters approaching match day. The peak speed achieved in training was approximately 93% of that recorded during matches.

Conclusion: Activity profiles vary by playing position, and there is a wide range of performances that can be considered normal with regards to high-speed running and sprinting, but less so for total distance and peak speed. The pyramid-like distribution of the external training load aligns with established training principles, facilitating rest, loading, and peaking before a match. Nevertheless, practitioners should recognize the potential gap in achieving maximum running speed during training.

Sammendrag

Bakgrunn: I takt med den økende populariteten og profesjonaliseringen av kvinnefotball er det et behov for ytterligere forskning for å styrke utviklingen og helsen til disse utøverne. For trenere utgjør arbeidskravsanalysen et viktig ledd i treningsplanleggingen. Ved å etablere normative data for et gitt prestasjonsnivå, kan trenere enkelt evaluere prestasjonen til sine egne utøvere i forhold til disse referansene. I tillegg, med den økende bruken av overvåkningsteknologi, blir referanseverdier for å identifisere avvik fra typiske prestasjoner stadig viktigere. Til slutt er det en betydelig mangel på forskning som har kartlagt den eksterne treningsbelastningen til kvinnelige fotballspillere i typiske treningssykluser og der kohortene består av flere lag. Med utgangspunkt i disse utfordringene var målet med avhandlingen å beskrive kamp- og treningsbelastningen til kvinnelige fotballspillere på toppnivå i Norge.

Metode: Vi samlet data fra fire lag i *Toppserien* over to sesonger. I artikkel I analyserte vi belastningen i kamp på posisjonsnivå, mens vi i artikkel II etablerte referanseverdier for uvanlige endringer i kampbelastning. Artikkel III beskrev den ytre belastningen i typiske treningssykluser og sammenlignet forskjellene i ytre belastning mellom startende og ikke-startende spillere.

Resultater: Spillerne tilbakela i gjennomsnitt mellom 8934 og 10131 meter totalt i løpet av en kamp, samt gjennomførte høyhastighetsløp på mellom 1054 og 1894 meter, sprintløp på mellom 227 og 530 meter, akselerasjonsløp på mellom 433 og 578 meter, deselerasjonsløp på mellom 305 og 493 meter, og oppnådde topphastigheter på mellom 27 og 29 km·h^{-1}. De største forskjellene i belastning ble observert mellom midtstoppere og kantposisjonene (backer og kant), der sistnevnte tilbakela større distanser med tanke på høyhastighetsløp og sprinter. Vi observerte også bare trivielle til små (Cohens D_z: 0.07-0.20) reduksjoner i aktivitet i periodene etter de mest intensive periodene sammenlignet med kampgjennomsnittet. Den observerte kamp-til-kamp-variabiliteten i høyhastighetsløpedistanse, sprintdistanse, og akselerasjon og deselerasjonsdistanse varierte fra 12 til 36%, mens topphastighet og total distanse varierte fra 4,5 til 5%. I lengre sykluser var det meste av belastningen konsentrert rundt midtuken, som en pyramide, med minimal differensiering i belastningen mellom startende og ikke-startende spillere imot kampdag. Toppfarten oppnådd i trening var omtrent 93% av det som ble registrert i kamp.

Konklusjon: Kampbelastning varierer basert på posisjon. Høyhastighetsløp og sprintdistanse viser stor variasjon fra kamp til kamp, mens totaldistanse og topphastighet varierer mye mindre. Den observerte pyramideformede belastningsfordelingen samsvarer med etablerte treningsprinsipper, der hvile, belastning, og topping prioriteres i forberedelsene til kamp. Likevel, trenere bør være oppmerksomme på den tilsynelatende mangelen på toppfart i treningshverdagen.

List of Tables

List of Figures

List of papers

Paper 1:

Winther, A. K., Baptista, I., Pedersen, S., Randers, M. B., Johansen, D., Krustrup, P., & Pettersen, S. A. (2022). Position specific physical performance and running intensity fluctuations in elite women's football. *Scandinavian Journal of Medicine & Science in Sports*, 32, 105-114.

Paper 2:

Baptista, I., Winther, A. K., Johansen, D., Randers, M. B., Pedersen, S., & Pettersen, S. A. (2022). The variability of physical match demands in elite women's football. *Science and Medicine in Football*, 6(5), 559-565.

Paper 3:

Winther, A.K., Baptista, I., Pedersen, S., Brito, J., Randers, M. B., Johansen, D., & Pettersen, S. A. (2023). An analysis of training load in highly trained female football players. *PLoS One (Accepted, in review)*.

Abbreviations

Acc_{dist} – Acceleration distance

$AccDec_{dist}$ – Combined acceleration and deceleration distance

CB – Center-backs

CM – Central midfielders

Dec_{dist} – Deceleration distance

FB – Full-backs

FW – Forwards

GPS – Global positioning system

GNSS – Global navigational satellite systems

HDOP – Horizontal dilution of precision

HSRD – High-speed running distance

RPE – Rate of perceived exertion

SpD – Sprint distance

SSG – Small-sided games

SWC – Smallest worthwhile change

TD – Total distance

VHSRD – Very high-speed running distance

WM – Wide midfielders

1 Introduction

1.1 Tracking technology and the training process

Sports training, as defined by, Viru & Viru (2000) involves systematically performing exercises to enhance physical abilities and acquire skills related to the technique in a sports event. While there is a wealth of literature pertaining to this subject, Impellizzeri et al. (2019) offer a clear framework for understanding the training process and how this relates to the design of the overall training plan. According to the authors, achieving specific performance adaptations requires targeting the systems underlying performance. Once training targets are set, manipulating training load becomes crucial for eliciting the desired response.

Drawing on the works of Coutts (2018), Impellizzeri (2019) describe training load as the input variable that is manipulated to elicit the desired training response. This classification of training load also depends on whether we are considering measurable aspects internally or externally to the athlete (Impellizzeri et al., 2019). External load, determined by the organization, quality, and quantity of exercises in the training plan, represents the prescribed physical work (Impellizzeri et al., 2019). In the context of football, external load is commonly monitored through tracking systems, with metrics such as total- and high-speed distance covered being prevalent indicators. Coaches use external load to tailor training, aiming to elicit specific physiological responses (Impellizzeri et al., 2019). This response corresponds to the internal training load, measured by indicators like heart rate and rate of perceived exertion (RPE). Reflecting the body's physiological response to the external load, internal load provides insight into the physiological strain experienced by the athlete during training (Impellizzeri et al., 2019).

It is through the external load construct that tracking technology emerges as a central aspect in the training process of football players. For example, Buchheit & Simpson (2017) identify three main objectives for integrating tracking technology into training: providing an objective assessment of external training load for sessions or matches, aiding in team-level programming of external training load, and facilitating decision-making for individual player training programs to enhance performance and prevent injuries (e.g., top-up training vs. unloading sequences, return-to-play progression). Torres-Ronda et al. (2022) further categorize these objectives into three overarching and overlapping purposes: Describing, Planning, and Monitoring. In the overlap of Describing~Monitoring, descriptive data is gathered by sport and/or position. This information is then utilized to plan physical outputs in the intersection of Describing~Planning before comparing the resulting physical outcomes with the training plan in the intersection of Planning~Monitoring. Consequently, the purposes described in Torres-Ronda et al. (2022) provide a strong rational for describing the training and match loads of female football players, as the normative data generated can be utilized by practitioners, particularly in the intersections of Describing~Monitoring and Describing~Planning.

1.1.1 GNSS-based tracking systems

A multitude of tracking systems are at the disposal of practitioners looking to monitor the external load of football players (Torres-Ronda et al., 2022). However, wearable devices that connect to one or several Global Navigational Satellite Systems (GNSS) – an umbrella term encompassing satellite constellations designed for positioning, navigation, and timing – rank among the most popular systems (Luteberget & Gilgien, 2020). Amid these, the USA-based Global Positioning System (GPS) stands out as the most widely integrated technology in commercially available tracking systems (J. J. Malone et al., 2017).

GNSS units operate by transferring data between the unit located on the athlete, and the available satellites orbiting the Earth (Luteberget & Gilgien, 2020). This information/data is transferred using a sampling frequency measured in Hertz (Hz), meaning, the higher the sampling frequency, the more information transferred per second (J. J. Malone et al., 2017). In the context of player monitoring, the speed measured by these systems usually serves as basis for calculating tracking metrics. Speed is usually ascertained through the Doppler shift method (Ellens et al., 2022), which involves measuring shifts in satellite signal frequency attributable to the movement of the receiver (Larsson, 2003). This receiver is typically worn on the upper back of the player affixed to a snug vest. From Doppler-derived speed, other common metrics used for player monitoring, such as distance and acceleration, can then be derived and further used to calculate distances and number and distances of runs in different intensity zones.

Currently, GNSS (or GPS) analysis remains the most effective and time-efficient for monitoring workload within the team sports. In addition, the validity and reliability of these systems are also improving with technological developments, such as advances in chipset technology and signal processing algorithms (Cardinale & Varley, 2017; J. J. Malone et al., 2017). Collectively, a large body of evidence underscores the validity and reliability, especially of 10Hz GNSS-based tracking devices, as a dependable method for measuring distance and speed in team sports (Beato et al., 2018; Rampinini et al., 2014; M. TU. Scott et al., 2016).

1.2 External training load metrics

As stated by Torres-Ronda et al. (2022), practitioners are besieged with a multitude of metrics from tracking systems. However, as recommended by Buchheit & Simpson (2017), if a tracking system is to significantly contribute to a training program's effectiveness, it is recommended to concentrate on variables that are simple to interpret whilst offering utility to practitioners. Moreover, these variables should demonstrate both validity and reliability to instill confidence when crucial decisions need to be made. When surveying the practices and perceptions of high-level football clubs, Akenhead et al. (2016) found that distances covered at various speeds and the occurrences of high-speed movement, accelerations and decelerations were the most commonly used metrics. Another study by Nosek et al. (2021) found that coaches were most interested in metrics describing “high intensity” actions and “work rate/intensity”. The same study also notes that players deemed feedback as positive to change their behavior, with total distance (TD), high-speed running (HSR) and sprint distances (SpD) as the information they would most likely act upon (Nosek et al., 2021). Furthermore, a survey by McCall et al. (2020) highlighted the importance of high-speed running, sprinting, and corresponding “worst-case scenarios” for preventing injuries. Taken together, these three studies provide valuable insights into what types of metrics sports scientists should prioritize if research is to translate into practice.

1.2.1 Total distance, high-speed running, and sprinting

All the metrics mentioned in the studies above correspond to several aspects of physical performance in football. While total distance does not differentiate between competitive levels (Choice et al., 2022; Mohr et al., 2008), it still serves as a valuable proxy for assessing overall training volume (Buchheit & Simpson, 2017). The volume of running at high speeds, however, seems to be of greater importance in terms of differentiating ability. For example, several studies have shown that players at higher competitive levels display greater volumes of high-speed running during match play compared to lower competitive levels (Andersson et al., 2010; Mohr et al., 2008). Though more recently, Scott et al. (2020) found few meaningful differences between international and domestic players. It is also interesting to note that the English Premier League, experienced a 12% increase in high-speed running- and a 15% increase in sprint distance from 2014/2015 to 2018/2019 (Allen et al., 2023), indicating that players are covering greater distances at higher speeds. Given the surge in competitiveness observed in the women’s game (de Araújo & Mießen, 2017), coupled with the paucity of scientific literature available (Kirkendall & Krustrup, 2022), one can speculate whether a similar, if not steeper, trend exists.

Straight-line sprinting is one of the rarest events in football yet is the most frequent action before goals for both the scoring and assisting player (Faude et al., 2012). Thus, while being an outlier in turns of events, a single sprint can have a disproportional effect on the outcome of a game. Beyond its impact on match dynamics, the importance of sprint ability also extends to the players’ overall physical performance during a game. As showed by Mendez-Vallanueva et al. (2011), the fastest players during sprint testing also reach the highest speeds during the game, and all players reach a high percentage of their maximal sprint speed regardless of position. Furthermore, the importance of high-speed running and sprinting is also underlined by its relation to injury. For example, both Buchheit et al. (2023) and Malone et al. (2018) found that male athletes who sprinted at high intensity (>95% of their maximal speed) during sport practice showed a lower risk of lower limb injuries than those who produced lower maximal speed (<85%). In another study, Buchheit et al. (2024) found that a weekly HSR-to-match ratio of 0.6 to 0.9, and a weekly SpD-to-match ratio of 0.6 to 1.1. was associated with reductions in injury occurrence. This has led to some authors suggesting that exposure to maximum speeds is a potential “vaccine” against hamstring injuries (Edouard et al., 2019), with exposure to high-speed football actions being suggested as a modifiable risk factor (Ekstrand et al., 2023).

Thus, the ability of total distance to gauge overall training volume, coupled with the association between high-speed running and sprinting, and game outcomes and injury prevention, underscore the importance of these metrics as integral parts in player monitoring.

1.2.2 Acceleration and deceleration

Another set of metrics that have gained more attention from sports practitioners and scientists in recent years are accelerations and decelerations. Several authors have noted that a significant part of the overall external load in team sports such as football are due to intense accelerations and decelerations (Harper et al., 2019; Vanrenterghem et al., 2017). These actions also impose distinct physiological and mechanical demands on the players (Harper et al., 2019; Vanrenterghem et al., 2017). For instance, Harper et al. (2019) write that accelerations incur a higher metabolic cost compared to decelerations, while decelerations result in a greater mechanical load (Dalen et al., 2016; Hader et al., 2016), often due to impactful peak forces and loading rates that can potentially cause more damage to soft-tissue structures if not efficiently mitigated (Harper & Kiely, 2018). The same authors also write that the frequency of high-intensity accelerations and decelerations during match play is commonly linked to declines in neuromuscular performance and signs of muscle damage post-match (De Hoyo et al., 2016; Gastin et al., 2019). Furthermore, Harper et al. (2019) also point to the differentiating ability of these metrics. For instance, elite athletes demonstrate a greater capacity to sustain a higher frequency and magnitude of accelerations and decelerations compared to their lower-performing counterparts (Draganidis et al., 2015). This capability could potentially contribute to improved match play performance, especially in situations requiring swift changes in velocity (Harper et al., 2019). Consequently, it is important that acceleration and deceleration can be appropriately quantified and monitored during training and competition to ensure athletes are adequately prepared for this load.

1.2.3 Peak periods

Practitioners also place value on metrics describing activity over shorter periods, particularly those related to high-speed running and sprinting (McCall et al., 2020). These types of metrics go by various names in the literature such as “peak periods”, “peak demands”, “peak locomotor demands”, “peak characteristics”, “duration-specific locomotor demands”, “maximal intensity periods”, or ‘worst-case scenarios” (Weaving et al., 2022). Regardless of nomenclature, the purpose is to capture the activity over a set period (e.g. 1-minute peak sprint distance, 2-minute peak sprint distance, etc.). Initially, these metrics were calculated by segmenting the data over predefined durations, often, per quarter or smaller predefined durations (e.g. 0 to 5, 5 to 10 minutes, etc.) (Weaving et al., 2022). However, as reported by Fereday et al. (2020), this approach underestimates 1–10-minute peak total distance and peak high-speed running distance by ∼7–10%, and ∼12–25% when compared to using a rolling average. This finding suggests that rolling averages may be a more suitable method of assessing peak periods in professional football (Fereday et al., 2020). Activity per minute during these periods also appears far greater compared to the match. For example, Riboli et al. (2022) demonstrated that the activity during 1-minute peak periods is 149%, 381%, 450%, and 781% of mean match activity in terms of total, high-speed, very-high speed, and sprint distance, respectively. This discrepancy is due to the match as a whole containing periods of inactivity, exacerbated when compared against high-speed metrics and shorter epochs (Weaving et al., 2022).

Compelling arguments for the use of peak periods in football practice are most notably articulated by Weaving et al. (2022). For instance, they reference a study by Barret et al. (2020) highlighting that position-specific-, possession-, small-sided game (SSG)-, tactical- and technical-based training modes account for 90% of total training drill prescription, with conditioning compromising the remaining 10%. According to Weaving et al. (2022), this finding underscores the need for strength and conditioning coaches to work within an environment where sport-specific practice compromises the bulk of training, and where the technical-tactical coach is in charge. For example, peak periods could serve as benchmark and evaluation tool during small-sided games when there is a greater emphasis on technical-tactical elements. Task constraints, such as field dimensions, could then be manipulated to achieve the desired external intensity (Riboli et al., 2020). Interestingly, the authors also note that general SSG prescription seems to underload players in terms of high-speed running but overload them in terms of mechanical work. This notion is based on a study by Lacome et al. (2018), comparing match peak periods of 1 to 5 minutes with those obtained under different SSG formats (4v4, 6v6, 8v8; all with goalkeepers). The authors also compare results from Gaudino et al. (2014), investigating activity during 4-minutes of 5v5, 7v7 and 10v10 SSG with those of Noak et al (2021), studying 3-minute peak periods in official matches. Both studies looked at Premier League players, making them comparable samples. The comparison showed that the relative total distance (range: 100.5-116.5 vs. 146-167 $m \cdot min^{-1}$) and relative high-speed running distance (range: 0.25-4.0 vs. 12.3-20.2 $m \cdot min^{-1}$) reported in Gaudino et al (2014) were much lower compared to the corresponding 3-minute peak periods reported in Novak et al. (2021). The highest values in Gaudino et al. (2014) were also far below the average intra-individual match-to-match variation (6.8% and 25.2%) reported in Novak et al. (2021), even when considering the lowest values for both metrics (TD: 136-156 $m \cdot min^{-1}$; HSRD: 9.2-15.4 $m \cdot min^{-1}$).

In summary, Weaving et al. (2022) suggest that peak periods can be valuable metrics for practitioners. This is because a) technical–tactical training modes constitute a significant portion of training, and b) technical–tactical coaches are primarily responsible for designing the task constraints in such training. As such, normative peak period data can be beneficial for practitioners in the evaluation and planning of external load in modes where there is a heightened emphasis on technical–tactical aspects and a reduced focus on physiological development.

1.3 Activity profiles of female football players

To reemphasize the framework by Torres-Ronda et al. (2022), tracking technology proves particularly valuable for acquiring descriptive data categorized by sport and position

(Describing~Monitoring). This data can then be utilized to plan physical outputs in the intersection of Describing~Planning before comparing the resulting physical outcomes with the training plan in the intersection of Planning~Monitoring. For practitioners working with female football players, descriptive data on match activity at different standards can therefore be a valuable resource, providing reference values or normative data that can help inform practice. For instance, authors such as Buchheit et al. (2024) have expressed the correlation with injury occurrence as a weekly training load to match ratio, while authors such as Baptista et al. (2020) and Stevens et al. (2017) have noted how scaling the session load based on the match load can provide context to the work performed. As an example, the sprint distance in a training session can be described both in terms of its absolute values (for example 250 meters), but also relative to the mean match sprint distance (50%, if the mean match sprint distance is 500 meters). In this way, the match load contextualizes the session in a practical way for practitioners.

1.3.1 Whole-match activity by position at the professional level

The most comprehensive systematic review of match-play characteristics in female football players is found in Harkness-Armstrong et al. (2022). Of the 69 studies included in their final selection, 39 presented whole-match absolute values of the most frequently reported physical characteristics (i.e. total distance, high-speed running distance, very-high-speed running distance, sprint distance, peak speed, and number of accelerations and decelerations).

Since activity increases the higher the playing standard, it is essential that normative data is created for each level. Activity is also likely dependent on player position; thus, one also needs to look at the interaction between playing standard and position. Diving deeper into the material of Harkness-Armstrong et al. (2022) one can find eight studies that fit these criteria (table 1). Of these studies, only four have measured physical performance using 10 Hz GPS, with thresholds differing between each study.

Table 1: Studies that have investigated the effect of position on physical performance in domestic and international level female football players.

Study[a]	Abbr.	Level	Metrics[bc]	Tracking system
Datson et al. (2017)	D	INT	[HSRD: 19.8-25.1, VHSRD: >19.8, SpD: >25.1]	25 Hz multi-camera
Mara et al. (2017)a	M	DOM	[ACC >2, DEC: < -2]	25 Hz optical
Mara et al. (2017)b	M	DOM	[HSRD: 12.24-19.44, SpD: >19.44]	25 Hz optical
Nakamura et al. (2017)	N	DOM	[SpD: >20]	5 Hz GPS
Nakamura et al. (2017)	N	DOM	[SpD: >19.37]	5 Hz GPS
Panduro et al. (2021)	P	DOM	[HSRD: >15, VHSRD: >18, SpD: >25, ACC: >3, DEC: < -3]	10 Hz GPS
Ramos et al. (2019)a	R	INT	[HSRD: 15.6-20, SpD: >20, ACC: >1 DEC: < -1]	10 Hz GPS
Scott et al. (2020)a	S	DOM	[HSRD: >12.5, VHSRD: >19, SpD: >22.5]	10 Hz GPS
Trewin et al. (2018)a	T	INT	[HSRD: >16.48, SpD: >19.98, ACC: >2.26]	10 Hz GPS

[a] Extracted from Harkness-Armstrong et al. (2022)

[b] HSRD and SpD in $km \cdot h^{-1}$

[c] ACC and DEC in $m \cdot s^{-2}$

From figure 1, one can see that players within these categories typically cover total distances of 9000-11000 m (A), HSRD of 500-3000 m (C), SpD of 20-850 m (D), and reach peak speeds of 27-31.5 $km{\cdot}h^{-1}$, while performing 7 to 475 accelerations, and 13 to 473 decelerations. Some of these ranges are quite broad, however, partly due to between-study differences in methodology, such as the tracking system and thresholds used.

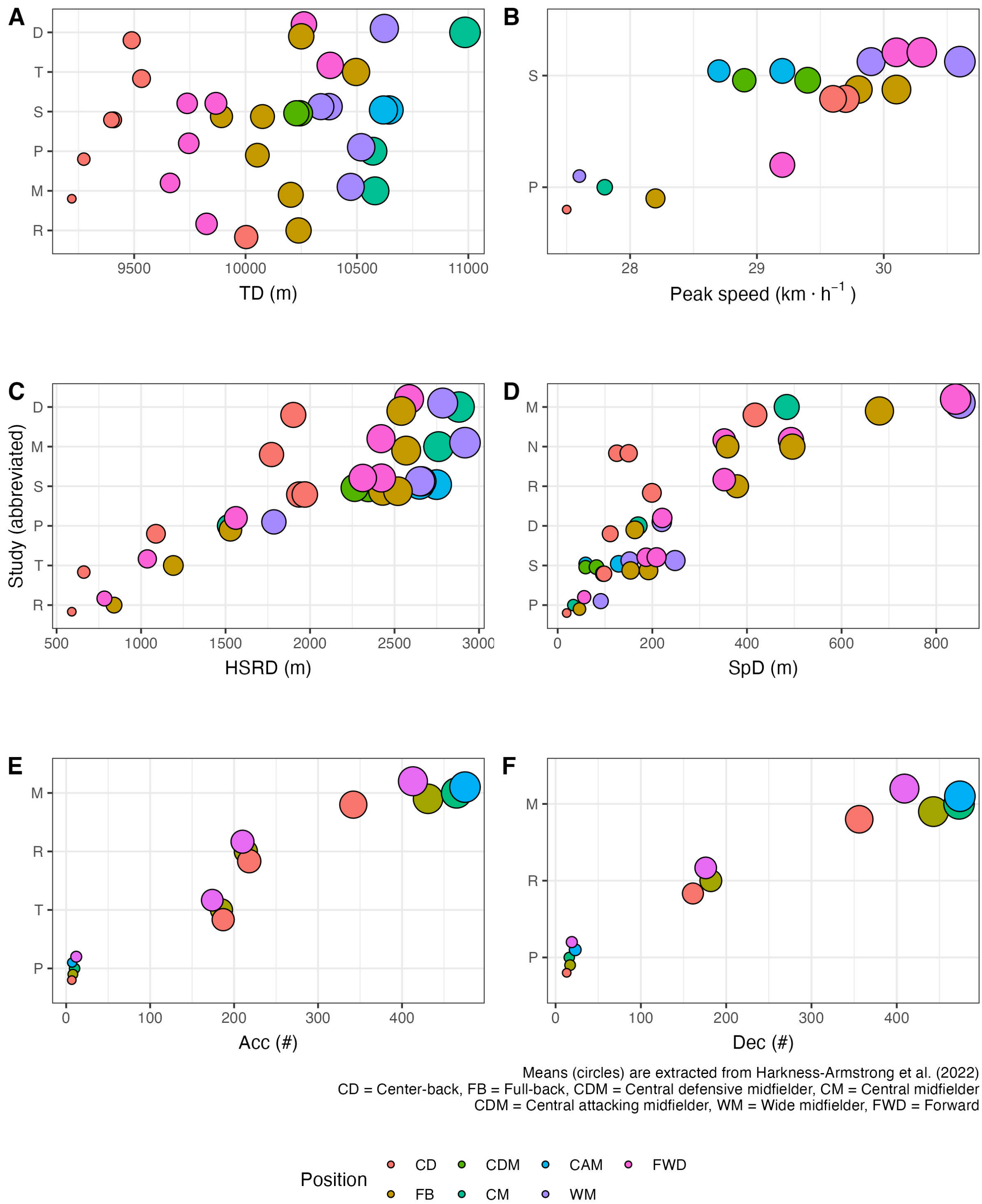

Figure 1: A) Total distance, B) peak speed, C) HSRD, D) SpD, E) accelerations, and F) decelerations in domestic and international level female football players.

1.3.2 Peak periods

Out of the studies reported in Harkness-Armstrong et al. (2022), only six reported peak periods in women's match play, with just one study applying a moving/rolling average analysis. As mentioned previously, fixed periods tend to underestimate 1–10-minute peak total distance and peak high-speed running distance approximately by 7–10%, and 12–25%, respectively, compared to using a rolling average (Fereday et al 2020).

Table 2: Studies in Harkness-Armstrong et al. (2022) that have described peak periods in domestic (DOM) and international (INT) level female football players.

Study[a]	Abbr.	Level	Metrics[bc]	Tracking system
Datson et al. (2017)	D	INT	[HSRD: 19.8-25.1, VHSRD: >19.8, SpD: >25.1]	25 Hz multi-camera
Mara et al. (2017)a	M	DOM	[ACC >2, DEC: < -2]	25 Hz optical
Mara et al. (2017)b	M	DOM	[HSRD: 12.24-19.44, SpD: >19.44]	25 Hz optical
Nakamura et al. (2017)	N	DOM	[SpD: >20]	5 Hz GPS
Nakamura et al. (2017)	N	DOM	[SpD: >19.37]	5 Hz GPS
Panduro et al. (2021)	P	DOM	[HSRD: >15, VHSRD: >18, SpD: >25, ACC: >3, DEC: < -3]	10 Hz GPS
Ramos et al. (2019)a	R	INT	[HSRD: 15.6-20, SpD: >20, ACC: >1 DEC: < -1]	10 Hz GPS
Scott et al. (2020)a	S	DOM	[HSRD: >12.5, VHSRD: >19, SpD: >22.5]	10 Hz GPS
Trewin et al. (2018)a	T	INT	[HSRD: >16.48, SpD: >19.98, ACC: >2.26]	10 Hz GPS

[a] Extracted from Harkness-Armstrong et al. (2022)

[b] HSRD and SpD in $km \cdot h^{-1}$

[c] ACC and DEC in $m \cdot s^{-2}$

Another interesting aspect is whether activity decreases following peak periods, which may indicate fatigue or pacing strategy (Bradley & Noakes, 2013). While both Andersson et al. (2010) and Trewin et al. (2018) presented data on post-5-minute periods, neither of them actually compared the activity during these periods to the mean match activity. One could hypothesize that a substantial decrease in post peak period activity relative to the mean match activity to be indicative of fatigue or pacing.

1.4 Variability and the smallest worthwhile change

Another concept that should be important to practitioners is the smallest worthwhile change (SWC). This concept was most notably expanded upon in an article by Hopkins et al. (1999), wherein the SWC was described as the magnitude of performance enhancement required to make a difference to the model-winning prospect of an elite athlete. According to Hopkins et al. (1999), there are two factors that are important when deciding upon the SWC: the variation in an athlete's performance between events (also known as within-athlete variation or variability), and the variation in performance between athletes in the same event (also known as between-athlete variation or variability). For example, it could appear that a small enhancement in performance would be worthwhile for one of these athletes because it would put that athlete ahead of all the others. However, one also needs to consider within-athlete variation because this produces slightly different outcomes each time, unrelated to any true performance enhancements. To put it in a practical sense, a performance enhancement much smaller than the within-athlete variation would obviously have no effect on an athlete's chances of winning an event, while an enhancement much greater than the within-athlete variation would guarantee the athlete first place (Hopkins et al., 1999). The enhancement that begins to make a difference to the athlete's chance of winning is somewhere between these two extremes. Between-athlete variation represents the true variation in ability between athletes (Hopkins et al. 1999). Between-athlete variation has an important effect on medal winning in individual sports because the greater the spread in ability relative to within-athlete variability, the greater the enhancement needed to lift an athlete to first place from a lower ranking (Hopkins et al. 1999).

It is harder to discern what is worthwhile change in performance for a female football player, as there is no clear relationship between physical performance and match outcome. Instead, Hopkins (2004) recommends using 0.2 multiplied by the between-athlete standard deviation or CV% as a general guideline for setting the SWC for team-sport athletes. This recommendation is based on the effect size statistic and the associated guidelines (small: 0.2-0.5, medium: 0.5-0.8, large: > 0.8) in Cohen (1988). Although Cohen (1988) outlined several variations of the effect size statistics dependent on study design, all are essentially dividing the difference or change in means by some form of uncertainty. By multiplying the between-athlete standard deviation by 0.2, as Hopkins (2004) recommends, you get the difference or change in means equivalent to an effect size of 0.2. In practical terms this is equivalent to an athlete moving from the 50^{th} to the 58^{th}, from 80^{th} to 85^{th}, or from 95^{th} to 97^{th} percentile rank in a sample or population of athletes (Hopkins, 2004).

An example of how the SWC can be used in combination with statistical concepts to provide practitioners with reference values for meaningful changes, is the article by Oliva-Lozano et al. (2021). They provide three measures that practitioners can use to interpret changes in individual players: the observed match-to-match variability, 80% and 90% limits of agreement (LoA), and values signifying statistically significant changes. Observed match-to-match variability captures the variation in a player's performance across successive matches. This measure is derived using linear mixed models, which enable the partitioning of variability into components attributed to factors such as the match itself and within-player variability (Malcata & Hopkins, 2014). Importantly, observed match-to-match variability focuses solely on components relevant to monitoring individual performance, excluding variations stemming from differences between players or between teams (Oliva-Lozano et al. 2021).

The next measure, limits of agreement (LoA), denotes the range in which an individual's change score would fall a certain percentage of the time in consecutive pairs of trials (Hopkins, 2000). For instance, in Oliva-Lozano et al. (2021) the 90% LoA for total distance is ± 1333, meaning that from one game to the next, a male player has a 1 in 10 chance of running 1333 meters longer or shorter in the second game compared to the first. Finally, Oliva-Lozano et al. (2021) also reports values that would flag a change as statistically significant while considering the observed match-to-match variability, the SWC, and alpha levels of 0.10 and 0.05.

In summary, by integrating concepts such as observed match-to-match variability, LoA, and statistical significance, practitioners can gain insights into unusual changes in individual players' performance metrics across successive matches.

1.5 In-season training

The rational for exploring the in-season training of female football players is explained by us in paper III but will be reiterated here for introductory purposes. In the beginning of the paper we write how variations in training load in football are commonly observed at the microcycle level, where external load is typically manipulated based on the number of days between matches (Clemente et al., 2014; Morgans et al., 2014). The “horizontal alternation” principle (Buchheit et al., 2018, 2021) is often mentioned in tandem, which suggests targeting specific physical capacities like strength, endurance, or speed on designated days. This approach aims to develop each capacity while minimizing physiological interferences (Buchheit et al., 2018; Fyfe et al., 2014), and is often applied within the "days before the match" (MD-) / "days after a match" (MD+) framework. For instance, with six days between matches, three "acquisition" days (MD-4, MD-3, and MD-2) might be interspersed with one or two "recovery" days (MD+1 to MD+2) and one "tapering" day (MD-1), dedicating each "acquisition" day to a specific capacity, aiming for overall development whilst allowing adequate recovery time (Buchheit et al., 2018).

Furthermore, few studies have explored the periodization of training load in female football teams. Karlsson et al. (2023) found a Norwegian team differentiating their training load in longer cycles (with 5-7 training days available), much in resembles to the horizontal alternation principle. On the other hand, Diaz-Seradilla et al. (2022) noted that with four days between matches, MD was more demanding than any training day, and all external training load variables were higher on MD-3 compared to other training days. Last, Romero-Moraleda et al. (2021) observed that in cycles with five days between matches, the match was the most demanding session, and MD-4 and MD-3 consistently produced the greatest physiological and biomechanical loads, with MD+1 showing the lowest values. In summary, while these papers have explored differences between training days in specific cycle lengths, limited information exists on training load across a broad range of cycles. Additionally, existing studies have focused solely on players with over 60 minutes of playing time, leaving gaps in our understanding of the training load for non-starters.

2 Aims of the thesis

The primary objective of this dissertation was to conduct a comprehensive analysis of the match activity and training loads of highly trained female football players, with the goal of offering practitioners valuable insights and practical implications in the training of female footballers. To achieve this, we specifically aimed to:

1. Describe the activity profiles of highly trained female football players by playing position and investigate fluctuations in activity following peak periods.
2. Establish reference values for unusual changes in metrics commonly used in player monitoring.
3. Track the microcycles of professional football teams and compare differences in external training load by day and by squad status (starter versus non-starter).

3 Methods

This section provides an overview of the data collection process for the entire project, as well as individual summaries of the statistical analysis conducted for each paper.

3.1 Ethical approval

Before commencing the project, we applied for ethical approval through the Regional Committee for Medical and Health Research Ethics - Northern Norway (reference number 53884) but were exempted since the data collection did not include a biobank, medical or health data related to illness, or interfered with the regular operation of the players. We also obtained approval from the Norwegian Centre for Research Data (reference number: 296155) before gathering written informed consent from the players in our studies. These players represented four teams in the Norwegian premier division, and one team in the Danish premier division, classified as highly trained according to the criteria outlined by McKay et al. (2022). Teams were pre-selected by the supervisors in cooperation with Toppfotball Kvinner (the organization responsible for organizing and overseeing women's elite football in Norway) based on existing knowledge of the competence and consistency surrounding the sporting apparatus of these teams. This decision was rooted in the belief that prioritizing a more consistent data collection was preferable to random selection.

3.2 Data sources

Starting in March 2020, a prospective observational study was conducted in which tracking data from training and matches over two full seasons were collected using STATSports Apex (Newry, Northern Ireland), with a sampling frequency of 10 Hz. The validity and level of accuracy (bias <5%) of this tracking system have been demonstrated by Beato et al. (2018). All teams trained and played home matches on artificial grass, with only occasional away games on natural grass. Training sessions usually started between 10 AM and 4 PM, with matches typically played between 1 PM and 9 PM during weekends. During training and matches, players wore their Apex unit on their upper back, adhering to manufacturer instructions (Figure 2). Furthermore, to minimize inter-device errors (Beato et al. 2018), each player used the same unit throughout data collection. Event data pertaining to fixtures and lineups were gathered via API-Football (2023) and NIFS (Norsk Internasjonal Fotballstatistikk) (2023) and merged with the tracking data during the processing of the files.

Figure 2: Placement of the STATSports Apex unit.

3.3 Data processing

All papers followed the reporting standards proposed by Malone et al. (2017) starting with raw GNSS data being exported from the manufacturer's software (STATSports Sonra, Newry, Northern Ireland) into a Python-based processing pipeline. The smoothing method, minimal effort duration, and thresholds used for calculating the metrics TD, HSRD, SpD, and accelerations and decelerations, are shown in Table 3, along with the most important independent variables used in each study. After processing, we performed statistical analysis on the aggregated data using the methods described in each paper, reiterated below.

Table 3: Smoothing, minimum effort duration (MED), thresholds, and independent variables in each study.

	Paper I & II	Paper III
Smoothing		1 second moving average
HSRD ($km \cdot h^{-1}$)	16	16
SpD ($km \cdot h^{-1}$)	20	20
Peak periods (minutes)	1 and 5; 1	
Acc/Dec ($m \cdot s^{-2}$)	2.26*	2.26
MED (s)	0.3	
Independent variables	Position, Team, Player ID, Match ID	MD (+-), Cycle, Squad Status, Player ID

*Raw acceleration calculated over 0.6 seconds

3.4 Statistical analysis

3.4.1 Paper I

This study included data from a single season compromising 60 matches, involving 108 female football players (22 ± 4 years of age) from the four Norwegian clubs participating in the project. We included only players who completed a minimum of two full-time matches and excluded goalkeepers and observations where players had less than 90 minutes of playing time. This resulted in an initial sample of 501 observations with 108 missing values. These values were subsequently omitted in the complete case analysis, resulting in a final of 393 match observations (M_{obs}) from 54 players. These players were categorized into different positions: center-backs (CB, n = 10, M_{obs} = 113), full-backs (FB, n = 11, M_{obs} = 84), central midfielders (CM, n = 16, M_{obs} = 105), wide midfielders (WM, n = 9, M_{obs} = 57) and forwards (FW, n = 8, M_{obs} = 34). The mean ± standard deviation of number of satellites and horizontal dilution of precision was 17.5 ± 2.8 and 1.4 ± 0.6, respectively.

To examine between positional differences in full match and peak metrics, we created a statistical model for each metric with *Position* as the fixed effect and *Team*, *Match ID× Team*

ID, and *Position × Player ID* as the random effects. These interaction terms were incorporated to address the nested structure of the data. For within-positional differences in peak, next, and mean periods we specified models with *Position*, *Period*, and the interaction term (*Position × Period*) along with the aforementioned random effects. We utilized the Tukey method to adjust for multiple comparisons, with an α-level set at 0.05 as the level of significance. Effect sizes were calculated using Cohen's D_z (Lakens, 2013). All statistical analyses were conducted using the lme4 (Bates et al., 2015) and emmeans4 (Lenth et al., 2018) packages.

3.4.2 Paper II

For this study we utilized the same dataset as for paper I, ensuring that sample characteristics remained consistent across the two papers. To decompose the various sources of variability (including between-team, between-position, between-player, between-match, and the residual within-player variability) and to provide reference values for interpreting changes in match physical performance, we followed the methodology outlined by Oliva-Lozano et al. (2021).

For each metric, we utilized a random effects model, specifying random intercepts for *Team*, *Position*, *Player ID*, and *Match ID*. Each random effect represented a source of variability and was expressed in raw units (as a standard deviation) by modelling the original data. Additionally, these effects were expressed as a percentage (CV%) by modelling the log-transformed data before back-transformation of each estimate, as proposed Hopkins (2017).

Like Oliva-Lozano et al. (2021), we utilized our estimates of variability to provide a framework for practitioners to interpret individual changes in match activity. Specifically, we calculated 80% and 90% LoA by multiplying the square root of 2 with the corresponding t-values from a t-distribution with infinite degrees of freedom and the observed between-match variability (e.g., the pooled between-match and within-player variability). Furthermore, "practical" or more correctly statistically significant changes associated with α-levels of 0.10 and 0.05 were calculated using the formula: $change = threshold + observed\ between - match\ variability \times \sqrt{2} \times t_{\alpha,\infty}$. Here, the observed between-match variability was the same as described above, while the threshold term was equivalent to the smallest worthwhile change (0.2 * the observed between-player variability – or the pooled between-player and within-player variability). As in paper I, all statistical analyses were conducted using the lme4 (Bates et al. 2015) and emmeans4 (Lenth et al. 2018) packages.

3.4.3 Paper III

This study included data from two seasons, involving 100 female football players (22.3 ± 3.7 years of age) from the same four teams as in the previous studies. Unlike the previous papers we put a stronger emphasis on handling missing data by following the recommendations by Bache-Mathiesen et al. (2022), Borg et al. (2022), and J.J Malone et al. (2017). As described in the paper, we set all the metrics as missing on sessions with a mean horizontal dilution of precision > 5 or a mean number of satellites <12. We also set peak speed as missing if above

32 km·h^{-1} based on theoretical max speed values of 29.2 ± 1.4 km·h^{-1} in a similar cohort (Haugen et al., 2020).

Furthermore, we described the initial dataset as including one observation for each squad player for each day throughout the competitive season (lasting 157 and 176 days in 2020 and 2021, respectively), totaling 12879 observations, with 7646 missing. We opted to remove all observations on MD+1 since it typically was a day off with a substantial amount of missing data (2208 out of 2426 observations). We also removed all observations in cycles with four training days due to too few observations (171 in total with 132 missing).

To impute missing data, we utilized multiple imputation with predicted mean matching, consistent with the recommendations by Bache-Mathiesen et al. (2022). Using the mice package (Van Buuren & Groothuis-Oudshoorn, 2011) in R, we applied the mice.impute.pmm function, including all dependent variables in addition to day number in the model, generating five imputed datasets for subsequent analysis.

In the statistical analysis, we modeled duration, TD, peak speed, and $AccDec_{dist}$ in R using the lme4 package (Bates et al. 2015), while HSRD and SpD were modelled in the same software using glmmTMB (Magnusson et al., 2017). All models included the interaction between *Match Day and cycle* and *Squad status* (*MDxCycle×SquadStatus*) as fixed effects and *Player ID* and *Team ID* as random effects. In addition, HSRD and SpD were modelled using the tweedie family with a log link function. We also examined, only for the starters, the differences in training load between each day within each cycle before comparing differences in training load between starters and non-starters within each day. Here, the package emmeans (Lenth et al. 2018) was used to compute estimated marginal means, using the Sidak method to adjust for multiple comparisons between the days and the Tukey method for pairwise comparison between starter and non-starters. We also conducted the same statistical analysis on the non-imputed dataset with only complete cases for sensitivity purposes.

4 Results

4.1 Paper I

4.1.1 Whole-match and peak activity by position

There were significant differences between playing positions across various metrics, highlighting a consistent pattern wherein players in wide positions covered more distance compared to those in central positions. Specifically, the results revealed that center-backs (CB) covered less distance than full-backs (FB) and central midfielders (CM) in terms of total distance (TD) and high-speed running distance (HSRD). Additionally, wide midfielders (WM) exhibited higher HSRD than CM and forwards (FW). Regarding SpD, CB covered less ground than FB, WM, and CM, while WM also surpassed FW in this aspect. Furthermore, WM outperformed FW in both TD and HSRD. Upon analyzing acceleration profiles, WM demonstrated higher acceleration distance (Acc_{dist}) than CB and CM, and higher deceleration distance (Dec_{dist}) than CB, CM, and FW (Figure 3 and 4). Peak speed, however, was not significantly different between any position.

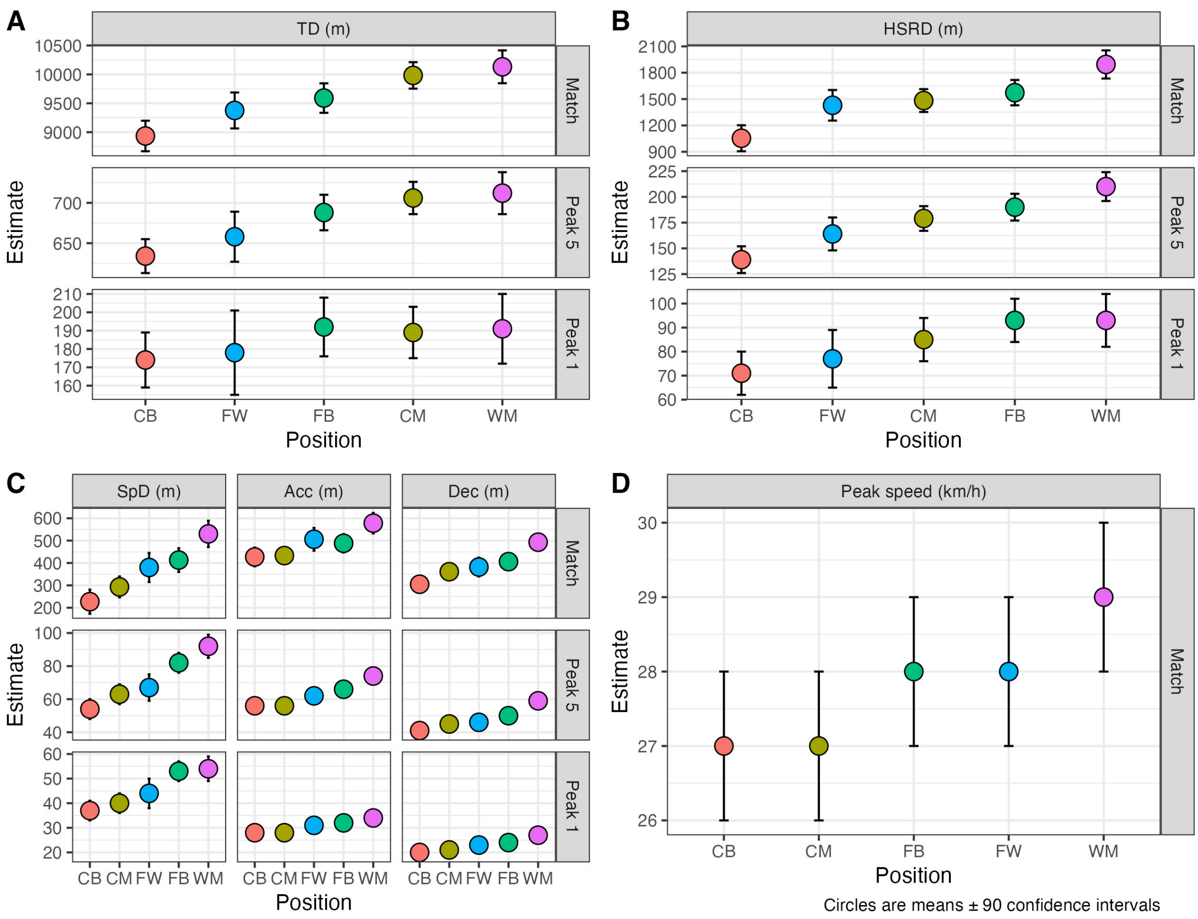

Figure 3: A) TD, B) HSRD, C) SPD, Acc, Dec, and D) peak speed, by playing position and period.

There were no significant differences between positions in 1-minute peak TD. However, in the case of peak 5-minute TD, three playing positions (FB, WM, and CM) displayed significantly higher values compared to CB (Figure 4). FB and WM also exhibited higher 1- and 5-minute peak HSRD than CB, with WM also surpassing CM and FW in the 5-minute peak (Figure 4). Similar trends were observed in terms of SpD. FB and WM had higher values in the 1-minute peak than CB and CM, and in the 5-minute peak than CB, CM, and FW. WM consistently had the highest values Acc_{dist} and Dec_{dist} in both 1- and 5-minute peak periods. Specifically, during the 1-minute peak, WM's results were significantly higher than CM, and during the 5-minute peak, they were higher than CB, CM, and FW.

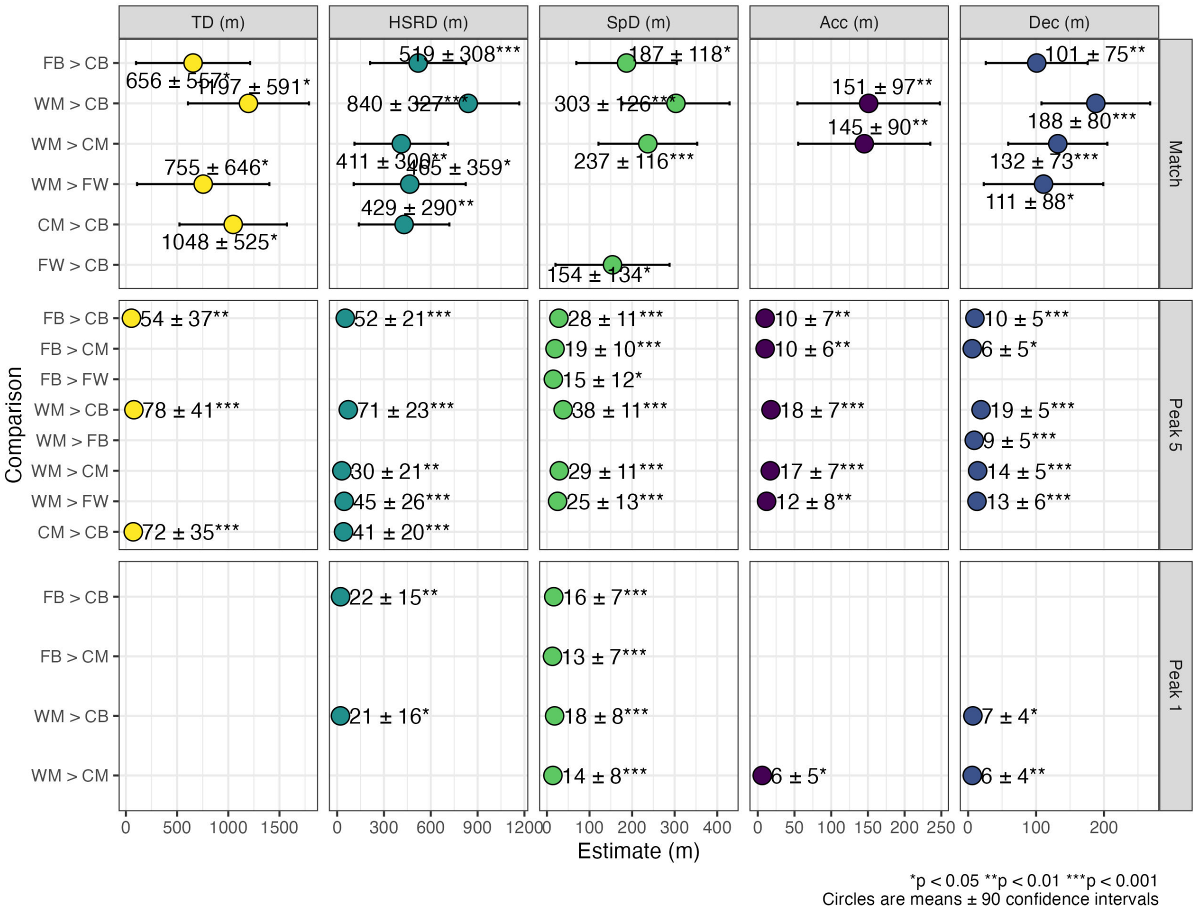

Figure 4: Statistically significant between-positional differences in TD, HSRD, SpD, Acc, and Dec for whole-match, peak 5-, and peak 1-minute metrics.

4.1.2 Activity following peak periods

In the examination of peak, next, and mean 5-minute periods, a consistent trend was observed for every playing position, as depicted in Figure 5. Regardless of the variable under analysis, the results consistently showed higher intensities during the peak 5-minute period compared to both the subsequent and mean 5-minute periods. In 1-minute peaks, both CB, FB, CM, and FW exhibited significantly higher values for HSRD, SpD, Acc_{dist}, and Dec_{dist} compared to the following 5-minute period (Figure 6). This trend was also observed for WM, except for HSRD, where the difference was not significant.

Furthermore, and regardless of peak period length, the next 5-minute periods consistently presented lower values, compared to the mean 5-minute values for each variable, however these differences were trivial to small, with effect sizes (Cohen's D_z) ranging from 0.07 to 0.20.

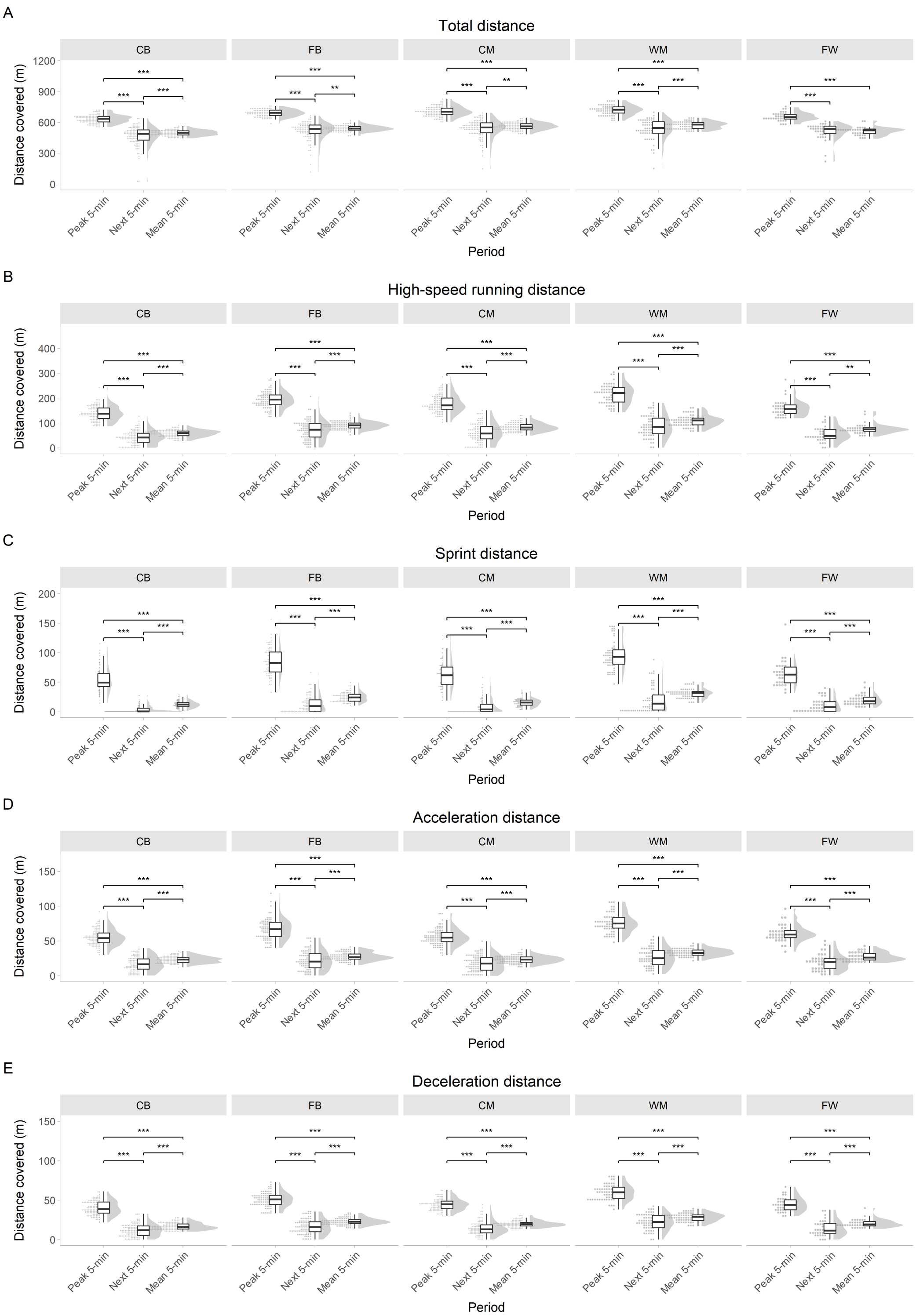

Figure 5: Distance covered during 5-minutes peaks, the next 5 minutes, and the mean 5-minute period, for total distance (A), high-speed distance (B), sprint distance (C), acceleration distance (D), and deceleration distance (E).

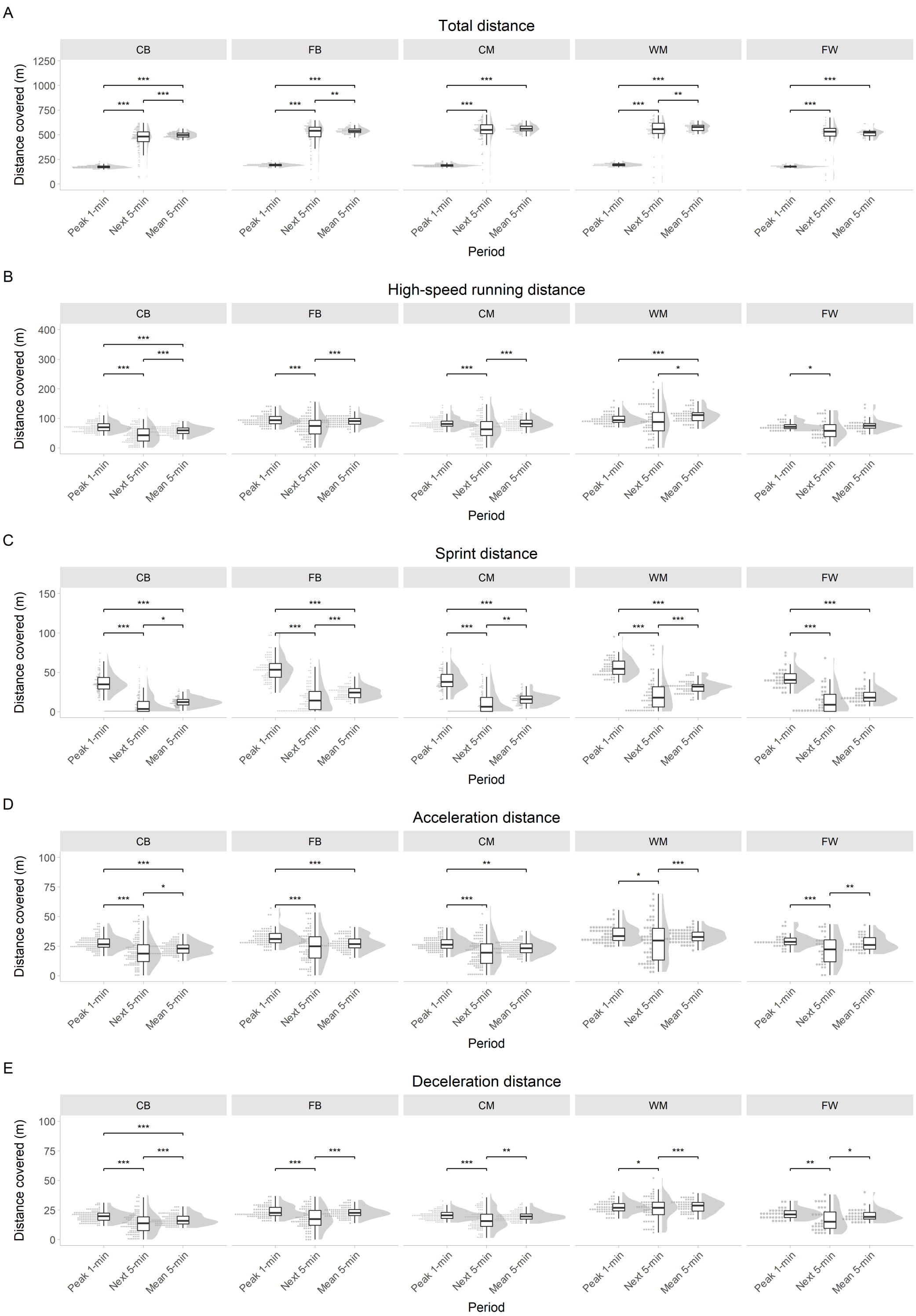

Figure 6: Distance covered during 1-minute peaks, the next 5 minutes, and the mean 5-minute period, for TD (A), HSRD (B), SpD (C), acceleration distance (D), and deceleration distance (E).

4.2 Paper II

4.2.1 Match-to-match variability

The decomposed variability of whole match and 1-min peak metrics are presented in Table 4. The variability between players (mean CV% range: 4-37%) and between positions (mean CV% range: 2-39%) was greater than that observed between teams, where there was minimal variability. Sprint distance saw the highest variability in each grouping factor, while variability in general was lower for peak compared to whole match metrics.

Table 4: Decomposed variability of whole-match and 1-minute peak metrics.

Measure	Metric	Period	Between-match	Between-player	Between-position	Between-team	Within-player/residual
SD	TD (m)	Match	335 (278 - 393)	473 (379 - 547)	456 (132 - 749)	37 (0 - 212)	259 (239 - 277)
SD	HSRD (m)	Match	132 (103 - 154)	272 (222 - 323)	288 (95 - 446)	51 (0 - 137)	160 (148 - 171)
SD	SpD (m)	Match	40 (29 - 49)	103 (84 - 122)	111 (31 - 172)	0 (0 - 39)	73 (68 - 78)
SD	AccDec (m)	Match	48 (35 - 59)	135 (105 - 156)	119 (37 - 204)	0 (0 - 51)	86 (80 - 91)
SD	AccDec (#)	Match	12 (10 - 15)	28 (23 - 33)	19 (0 - 32)	0 (0 - 11)	20 (18 - 21)
SD	Peak speed (km/h)	Match	0.3 (0.0 - 0.4)	1.2 (0.9 - 1.4)	0.5 (0.0 - 0.9)	0.0 (0.0 - 0.5)	1.2 (1.1 - 1.3)
SD	TD (m)	PP1	6 (4 - 7)	7 (5 - 8)	8 (2 - 14)	2 (0 - 4)	10 (10 - 11)
SD	HSRD (m)	PP1	7 (5 - 8)	8 (6 - 10)	10 (2 - 15)	2 (0 - 6)	13 (12 - 14)
SD	SpD (m)	PP1	3 (0 - 4)	6 (4 - 7)	8 (3 - 12)	1 (0 - 3)	11 (10 - 12)
SD	AccDec (m)	PP1	1 (0 - 2)	5 (4 - 7)	5 (0 - 7)	0 (0 - 2)	7 (6 - 7)
SD	AccDec (#)	PP1	0 (0 - 0)	1 (1 - 1)	1 (0 - 1)	0 (0 - 0)	1 (1 - 1)
CV%	TD (m)	Match	4 (3 - 4)	5 (4 - 6)	5 (2 - 8)	0 (0 - 2)	3 (3 - 3)
CV%	HSRD (m)	Match	10 (8 - 12)	19 (15 - 23)	23 (7 - 37)	1 (0 - 7)	12 (11 - 13)
CV%	SpD (m)	Match	14 (10 - 18)	37 (28 - 46)	39 (9 - 67)	0 (0 - 13)	28 (26 - 30)
CV%	AccDec (m)	Match	6 (5 - 7)	17 (13 - 20)	14 (3 - 25)	0 (0 - 7)	10 (10 - 11)
CV%	AccDec (#)	Match	6 (5 - 8)	14 (11 - 17)	9 (0 - 16)	0 (0 - 6)	10 (9 - 10)
CV%	Peak speed (km/h)	Match	1 (0 - 2)	4 (3 - 5)	2 (0 - 4)	0 (0 - 2)	4 (4 - 5)
CV%	TD (m)	PP1	3 (2 - 4)	4 (3 - 5)	5 (2 - 7)	1 (0 - 2)	6 (5 - 6)
CV%	HSRD (m)	PP1	8 (5 - 10)	11 (8 - 13)	13 (3 - 21)	1 (0 - 5)	17 (16 - 18)
CV%	SpD (m)	PP1	6 (0 - 10)	15 (10 - 19)	20 (7 - 32)	0 (0 - 7)	28 (26 - 31)
CV%	AccDec (m)	PP1	2 (0 - 4)	11 (8 - 13)	9 (1 - 15)	0 (0 - 4)	13 (13 - 14)
CV%	AccDec (#)	PP1	2 (0 - 4)	7 (5 - 9)	7 (1 - 11)	0 (0 - 3)	13 (12 - 13)

SWCs, observed match-to-match variability, and values indicating statistically significant changes activity are summarized in Table 5. Except for SpD (29.4 vs. 31.9%), all other metrics exhibited higher observed match-to-match variability in 1-minute peaks compared to the match (6.5 vs. 4.6%; 18.7% vs. 15.9%; 12.9 vs. 11.7%; for TD, HSRD and AccDec, respectively). Based on these findings, changes of ±9% ($\alpha = 0.10$) and ±12% ($\alpha = 0.05$) in whole match metrics of TD and peak speed are necessary to flag a change as unusual. For HSRD (33%; 42%), SpD (68%; 84%) and AccDec (25%; 31%) these thresholds ($\alpha = 0.10$; $\alpha = 0.05$; respectively) are notably higher. Except for SpD, relatively larger changes are required in 1-min peaks compared to whole match values to reach the same α-level.

Table 5: Smallest worthwhile change (SWC), observed match-to-match variability, and values indicating statistically significant changes within players between matches.

Metric	Period	Unit	SWC BP	SWC OBP	Observed between-match	LoA80	LoA90	p = 0.10	p = 0.05
TD	Match	m	95	108	423 (379 - 470)	767	985	875	1093
HSRD	Match	m	54	63	207 (188 - 226)	376	482	439	545
SpD	Match	m	21	25	83 (77 - 89)	151	193	176	219
AccDec	Match	m	27	32	98 (91 - 106)	178	229	210	261
AccDec	Match	#	6	7	23 (21 - 25)	42	54	49	61
Peak speed	Match	km/h	0.2	0.3	1.2 (1.1 - 1.3)	2.2	2.8	2.5	3.1
TD	PP1	m	1	2	12 (11 - 13)	21	27	24	30
HSRD	PP1	m	2	3	15 (14 - 16)	27	34	30	38
SpD	PP1	m	1	2	11 (11 - 12)	20	26	23	29
AccDec	PP1	m	1	2	7 (6 - 7)	12	15	14	17
AccDec	PP1	#	0.1	0.2	1.3 (1.2 - 1.4)	2.3	3	2.6	3.3
TD (m)	Match	%	1	1.1	5 (4 - 5)	8.3	10.6	9.4	11.8
HSRD (m)	Match	%	3.8	4.6	16 (14 - 18)	28.8	36.9	33.3	41.5
SpD (m)	Match	%	7.4	9.8	32 (29 - 35)	57.7	74.1	67.5	83.9
AccDec (m)	Match	%	3.4	4	12 (11 - 13)	22.1	28.3	26.1	32.3
AccDec (#)	Match	%	2.8	3.5	12 (11 - 13)	21.2	27.2	24.7	30.7
Peak speed (km/h)	Match	%	0.9	0.9	4.5 (4.2 - 4.8)	8.1	10.4	9.4	11.7
TD (m)	PP1	%	0.7	1.4	7 (6 - 7)	11.8	15.1	13.2	16.5
HSRD (m)	PP1	%	2.1	4.1	19 (17 - 20)	33.9	43.5	38	47.6
SpD (m)	PP1	%	3	6.6	29 (27 - 32)	53.3	68.4	59.9	75
AccDec (m)	PP1	%	2.2	3.6	14 (13 - 15)	24.7	31.6	28.2	35.2
AccDec (#)	PP1	%	1.5	3	13 (12 - 14)	23.4	30	26.3	33

BP = Between-player

OBP = Observed between-player

p = Change ± required to be statistically significant at alpha level

4.3 Paper III

4.3.1 In-season training

4.3.1.1 Match vs. training

Starters exhibited significantly higher values ($p < 0.001$) for TD, HSRD, SpD, $AccDec_{dist}$, and peak speed on MD compared to any training session independent of cycle length. The duration of MD was approximately 88 ± 1 minutes (min), shorter (7 ± 4 to 18 ± 4 min, $p < 0.001$) than training on most acquisition days (MD-5 to MD-3) in cycles with 5-7 days between matches.

4.3.1.2 Three days between matches

With three days available (1280 observations), there were no significant differences in duration and TD between MD+2 and MD-1. However, $AccDec_{dist}$, HSRD, SpD, and peak speed were slightly higher on MD-1 compared to MD+2, with differences of 108 ± 91 meters (m) ($p = 0.005$), 77 ± 38 m ($p < 0.001$), 21 ± 13 ($p < 0.001$), and 2.2 ± 1.2 $km{\cdot}h^{-1}$ ($p = 0.01$), respectively.

4.3.1.3 Five days between matches

In cycles with five days between matches (2192 observations), TD, HSRD, SpD, $AccDec_{dist}$, and peak speed were all lower on MD+2 compared to the other training days, except for TD (81 ± 493 m, $p = 1.000$) and $AccDec_{dist}$ (47 ± 115 m, $p = 1.000$) on MD-1. Differences in TD and mean peak speed ranged from 2728 ± 434 to 1005 ± 597 m, and from 2.9 ± 1.9 to 5.0 ± 1.5 $km{\cdot}h^{-1}$, respectively, while differences in HSRD and SpD ranged from 82 ± 49 to 356 ± 74 m and from 24 ± 13 to 108 ± 30 m. Differences in $AccDec_{dist}$ ranged from 668 ± 122 to 251 ± 121 m. All metrics displayed greater values ($p < 0.001$) on MD-3 compared to the other days of the cycle, with the largest differences observed when compared to MD+2 and MD-1, respectively.

4.3.1.4 Six days between matches

In six-day cycles (3002 observations), all metrics showed greater values on MD-4 to MD-2 compared to MD+2 ($p < 0.001$). Similarly, both TD (ranging from 1712 ± 430 to 3087 ± 380 m), HSRD (82 ± 42 to 353 ± 111 m), SpD (20 ± 15 to 102 ± 27), and $AccDec_{dist}$ (320 ± 137 to 721 ± 110 m) were greater on MD-4 to MD-2 compared to MD-1. However, statistically non-significant differences in peak speed (0.7 ± 0.8 $km{\cdot}h^{-1}$, $p = 0.158$) were found between MD-4 and MD-1. Furthermore, MD-3 saw prolonged duration (11 ± 5 min, $p < 0.001$) and higher peak speeds (1.5 ± 1.0 $km{\cdot}h^{-1}$, $p < 0.001$) compared to MD-4, and greater TD (1323 ± 370 and 1374 ± 463 m, $p \leq 0.001$), HSRD (245 ± 65 and 188 ± 84 m, $p < 0.001$), SpD (81 ± 26 and 58 ± 29 m, $p < 0.001$), and $AccDec_{dist}$ (244 ± 103 and 408 ± 182 m, $p < 0.001$) compared to both MD-4 and MD-2. The only difference between MD-4 and MD-2 was in $AccDec_{dist}$ (157 ± 156 m, $p = 0.047$) and peak speed (1.1 ± 1.1 $km{\cdot}h^{-1}$, $p = 0.047$), with greater $AccDec_{dist}$ covered on MD-4, and higher peak speed on MD-2.

4.3.1.5 Seven days between matches

Seven-day cycles (1872 observations) showed a similar pattern to five- and six-day cycles, with all variables being greater on MD-5 to MD-3 compared to MD+2. There were also differences in the tapering stage of the cycle, with extended (11 ± 10 min, $p \leq 0.015$) practice time on MD-2 compared to MD-1, coupled with more TD (1021 ± 602 m, $p < 0.001$) and $AccDec_{dist}$ (162 ± 142 m, $p = 0.009$) covered. TD, HSRD, SpD, and $AccDec_{dist}$ were higher on MD-4 than any other training day. $AccDec_{dist}$ was greater on MD-5 versus MD-3 (211 ± 171 m, $p = 0.004$).

4.3.1.6 Starters vs. non-starters

Starters vs. non-starters displayed mostly small and non-significant differences in external training load, except on MD+2. Non-starters trained longer (7 ± 5 to 13 ± 4 min, $p \leq 0.001$) in cycles with 3-6 days between matches, resulting in more TD (731 ± 246 to 1197 ± 218 m, $p < 0.001$), $AccDec_{dist}$ (176 ± 68 to 346 ± 106 m, $p < 0.001$), HSRD (28 ± 23 to 51 ± 26 m, $p \leq 0.019$), and higher peak speeds (1.2 ± 1.2 to 1.7 ± 0.7 km·h^{-1}) on those days compared to starters.

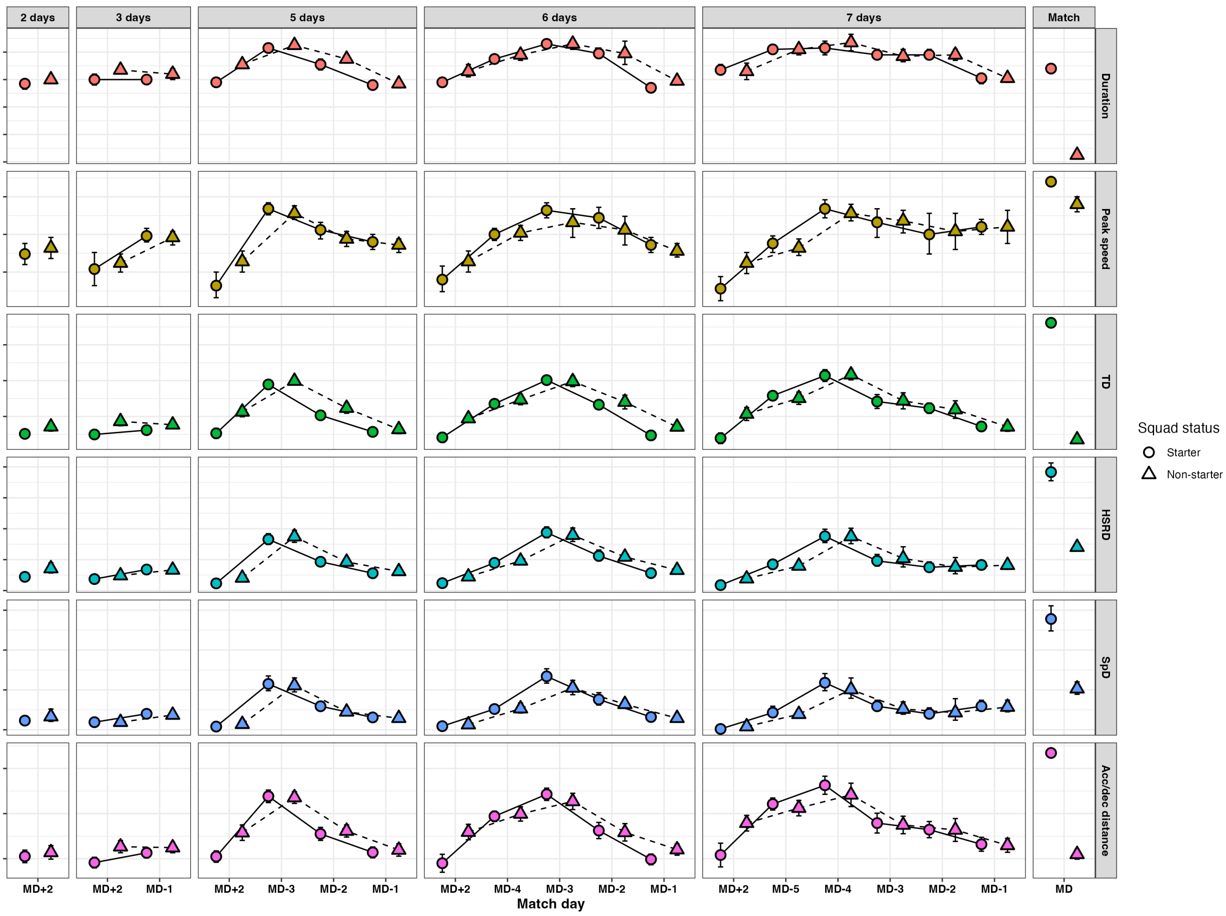

Figure 7: External training load by match day and cycle length.

5 Discussion

This chapter provides a concise overview of the main results from each paper, along with additional insights and perspectives that extend beyond the scope of individual papers.

5.1 Activity profiles by position

Paper I presented novel data on activity profiles from a large cohort of highly trained Norwegian female football players, along with fluctuations in activity following peak periods. Our findings align with previous studies (Figure 1), indicating that CD covers the least amount of distance irrespective of metric, while FB and WM cover the most HSRD and SpD. This discrepancy may stem from distinct positional responsibilities. For instance, FB and WM are likely tasked with maintaining width whilst contributing both offensively and defensively, while CD and FW are more likely to adopt stationary roles (Ju et al., 2023).

Greater distances were covered in peak periods relative to subsequent periods and the overall match. However, although subsequent periods generally exhibited statistically lower activity levels compared to the mean match periods, these differences were negligible or minimal at best. Figures 5 and 6 also illustrate the wide range of responses observed in the subsequent periods, indicating very large fluctuations in activity levels around the corresponding mean. These figures corroborate the findings of Trewin et al. (2018), who reported match-to-match variations in post 5-minute peak metrics ranging from 14 to 262%, with higher thresholds (such as SpD) associated with greater variability in response. Overall, this suggests that activity levels following peak periods tend to return to mean match level or lower but can vary considerably.

Furthermore, the overall output during each peak period may be considered low. For instance, the mean 5-minute- and mean 1-minute TD covered by a WM were 712 and 191 meters, respectively, translating to average speeds 8.5 $km \cdot h^{-1}$ and 11.46 $km \cdot h^{-1}$ during these periods. Additionally, there was a relative decrease in distance covered from peak 1-minute to peak 5-minute periods, consistent with findings by Riboli et al. (2023), who demonstrated an exponential decrease in distance covered as epoch length increases (see Figure 8).

Regarding the methods, one limitation is our decision to exclude observations under 90 minutes, which constrained our sample size unnecessarily. The better approach would have been to include all observations and control for the minutes played by each player, along with whether they were a starter or substitute. This would likely have reduced the confidence intervals around our estimates further, thus providing more precise activity profiles.

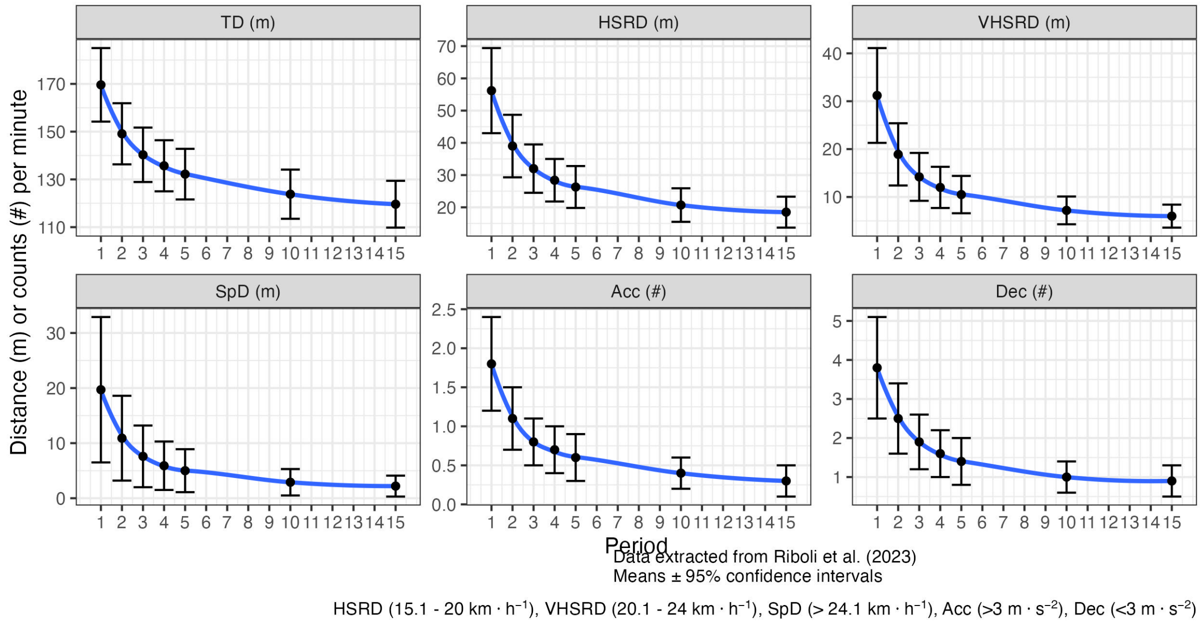

Figure 8: Decrements in distance- or counts per minute by epoch length.

The discussion section in paper I contains some conjecture that warrants clarification. One example is our assertion that *"preparing players to cope with the 5-min peak periods of the match do not necessarily mean that these players will be ready for the most demanding 1-min peaks, since the demands of 5-min peaks are not evenly distributed across every minute"*. This is a convoluted way of saying that there could be a substantial amount of low activity in 5-minute peak periods, potentially masking any instances of higher activity. Conversely, the shorter epoch lengths of 1-minute peak periods makes it less likely for instances of higher activity to be obscured.

We also conjecture on presumed demands and fatigue solely based on the distances of the peak periods, which in hindsight seems somewhat misguided. Making definitive statements on these aspects would require insights into the players' internal load, which we did not include in our study. Additionally, while a sprint effort in the 15th minute may register as "the peak period" it is probably the sprint in the 90th minute, when glycogen is depleted, that truly represents the peak demand, despite lower output. Conversely, it would not be inaccurate to suggest that the peak periods could contribute towards fatigue. For instance, Krustrup et al. (2022) observed a marked decrease in muscle glycogen (318 ± 105 and 248 ± 101 mmol·kg d.w^{-1}) and a marked increase in muscle lactate (14.3 ± 4.6 and 9.8 ± 3.7 mmol·kg d.w^{-1}) following intense periods in the first and second halves when compared to baseline (pre-match glycogen: 409 ± 62 mmol·kg d.w^{-1}, pre-match muscle lactate: 6.4 ± 3.7 mmol·kg d.w^{-1}). Additionally, repeated sprint ability was notably reduced following these periods, with moderate correlations to reductions in muscle glycogen and elevations in lactate concentration (Krustrup et al. 2022). While these periods were vaguely defined by the authors as "*as a period with a high number of high-intensity runs and sprints, evaluated by a UEFA Pro-License coach*", the authors note that intense periods could be related to peak periods. As indicated by the authors, their findings underscore that anaerobic energy production is markedly elevated during intense periods in women's match play (Krustrup et al. 2022).

While our reported distances for each peak period may seem low, it is important to consider that these distances can be accumulated in highly intensive ways. For instance, despite the mean peak 1-minute SpD for a WM being only 54 meters, various concurrent activities such as changes of direction, accelerations, and decelerations are likely taking place in this short span of time (Ade et al., 2016). Our suggestion that "*…the high intensity in the SpD 1-min peak period adds support to the prescription of speed endurance activities during training"*, is based on these assumptions, and, assuming the corresponding internal load aligns with Krustrup et al. (2022), would indicate that speed endurance training could effectively prepare players for these types of periods. Such training typically involves 20-90 seconds all-out bouts interspersed with recovery periods of 40 to 120 seconds, depending on the modality (Iaia & Bangsbo, 2010). Here, male players typically cover 125 to 131 meters of total distance in position-specific 30-second bouts, with very high-speed running (19.7 to 25.2 $km \cdot h^{-1}$) distances of 38 to 57 meters, combined with high lactate levels (~18 $mmol \cdot L^{-1}$) (Ade et al., 2021). This has further been shown to improve intense and repeated high-intensity exercise (Iaia & Bangsbo, 2010).

The above discussion highlights an important point for future studies, as more studies should try to explore how distance is accumulated within these periods in combination with internal load. These potential findings could serve as a basis for setting up position-specific courses.

5.2 Reference values for use in monitoring

Paper II presented reference values for three measures that practitioners can utilize to flag unusual or surprising changes in match performance: the observed match-to-match variability, limits of agreement, and the change required to achieve statistical significance. Additionally, this dissertation also adds the smallest worthwhile change (SWC) based on pure between-player variability and the SWC based on observed between-player variability to table 4, alongside raw scores.

It is noteworthy that the results for the observed match-to-match variability closely align with those reported in Trewin et al. (2018). For instance, Trewin et al. (2018) reported CVs of 6.4%, 33%, 53%, and 16% for TD, HSRD, number of sprints, and number of accelerations, respectively; while in our study, we found comparable figures of 5%, 16%, 32%, and 12% using similar metrics. Furthermore, the SWCs in Trewin et al. (2018) exhibit similarity to the SWC derived from the observed between-player in our study (1.8%, 7.5%, 9.4%, and 3.8% versus 1.1%, 4.6%, 9.8%, and 3.5%). Any disparities between our studies are likely attributable to differences in methodology and/or the fact that Trewin et al. (2018) sampled data over a more extended period (five years versus one season) which could account for the higher estimates.

The reported observed match-to-match variability has implications for precision in future studies or when monitoring players (Hopkins, 1997). For example, practitioners or researchers might be interested in comparing changes in match activity between two competition periods, such as the difference before and after a mid-season break, or after changes in coaching staff – scenarios common in modern football. In such cases the investigators aim to estimate the mean change between periods along with the uncertainty surrounding this estimate. When planning such studies, the sample size or number of measurements required can be determined based on the desired confidence interval (Rothman & Greenland, 2018). For crossovers or simple experiments without a control group, a 95% confidence interval can be calculated using the formula: 95% $CI = \bar{x} \pm \sqrt{2} \times t_{0.975,df} \times \frac{SD}{\sqrt{n}}$, where $\bar{x}$ is the sample mean, SD is the standard deviation, n is the sample size, t is the value of the t statistic for cumulative probability of 0.975, and df is the degrees of freedom (equal to n-1) (Hopkins, 1997). By isolating the margin error of error (MOE) ($MOE = \sqrt{2} \times t_{0.975,df} \times \frac{SD}{\sqrt{n}}$) and rearranging the formula to solve for n, we obtain $n = \frac{2t_{0.975,df}{}^2 SD^2}{MOE^2}$. Since the investigators are interested in changes in match activity, the observed match-to-match variability can be used as the SD in the formula. If investigators consider a margin of error equivalent to the observed match-to-match variability appropriate, and approximate $t_{0.975,\,df}$ to 2, the estimated sample size is eight change scores. For greater precision, if the margin of error is desired to be half of the observed match-to-match variability, the estimated sample size would be 32 changes scores. Consequently, metrics such as TD and peak speed are more reliable for detecting changes over shorter periods of time compared to HSRD and SpD, due to their lower CV%.

In hindsight, there are several methodological considerations that could have strengthened the paper. For instance, just like in paper I, the inclusion of minutes played as a fixed effect in our models would have increased the sample size considerably. Additionally, we specified position as a random effect when it just as well could have been specified as a fixed effect. Fixed effects, as explained by Malcata (2014), are predictors that affect the entire population in the same way, with all possible levels represented in the data. On the other hand, random effects are predictors whose levels are a random sample of the population and are often used to account for repeated measurements or clustering within each level of the random effect (Malcata 2014). In our context, position likely qualifies as a fixed effect since the demands and role responsibilities of a particular playing position likely affects each player similarly. Also, all levels of playing position are present in our sample (CD, FB, CM, WM, FW), and even if we operationalized this factor differently (for example, by splitting CM further into CAM, CM, and CDM) we still would have all levels represented.

We also chose to utilize the observed between-player variability to calculate the SWC, in contrast to the approach of Oliva-Lozano et al. (2021) who used the pure between-player variability. This decision resulted in a slightly larger threshold for determining significance, where for example a player transitioning from the 50^{th} to the 58^{th} percentile considers both differences between players and within-player variability. Furthermore, like Oliva-Lozano et al. (2021), we denote the change required to be statistically significant as “practically significant”, which should be clarified. According to Greenland et al. (2016), the p-value can be viewed as a continuous measure of the compatibility between the data and the entire model used to compute it, ranging from 0 for complete incompatibility to 1 for perfect compatibility. The smaller the p-value, the more unusual the data would be if every single model assumption (such that the effect is 0) were correct (Greenland et al. 2016). Extending this concept to the “practically significant changes” would mean that a change equal to or more extreme than this is unusual given the observed match-to-match variability and the SWC. Though whether this “unusualness” translates into practicality is unknown. As such the term “statistically significant changes” is a better term compared to “practically significant changes”.

In line with previous studies on variability in competitive performance (Malcata & Hopkins, 2014), we referred to the residual term in our models as “within-subject” or “within-player variability”. While this terminology is common in within-subject designs (Weir, 2005), it is important to recognize that the residual may encompass not only the true within-subject variability but also uncontrolled sources of variability. If we had included factors such as match date, minutes played, team and opposition Elo rating (originally used in Chess and subsequently as the basis for FIFA ratings (Hvattum & Arntzen, 2010)), and competition type in our models, the residual term may have been lower and more reflective of the true within-player variability.

Moreover, we could have added the intraclass correlation coefficient (ICC) to our results, which would have two main advantages. First, the ICC provides insights into to the proportion of variance explained by between-group difference (Weir, 2005), offering a measure of how much each random effect contributes to the overall variability. Second, the ICC also serves as a measure of repeatability (Malcata & Hopkins, 2014), and would have allowed us to assess whether the groups in each random effect consistently perform differently from each other.

5.3 In-season training

Paper III presented novel data on the external training load across various cycle lengths in highly trained female footballers along with differences between starters and non-starters. Two key findings emerged from our study. Firstly, teams adjusted their external load based on the number of days between matches, concentrating most of the training load towards the mid-week, succeeding, and preceding days of lower loads. Secondly, there was minimal differentiation in training load between starters and non-starters after MD+2 and onwards, regardless of cycle length. Our data suggests that teams typically incur the highest combined external load around three to four days before a match. This period of higher load precedes and succeeds days of lower load, aligning with common periodization principles. However, in shorter cycles with only two or three days between matches, most of the time is spent at lower loads, awaiting a mid-week game. Notably, there was less difference in $AccDec_{dist}$ compared to SpD on these days, possibly due to preferences for small-sided games.

Regarding training day differentiation, we observed no significant differences in SpD between MD-4 and MD-2, nor between MD-5 and MD-3 in longer cycles. However, there were significantly more $AccDec_{dist}$ covered on MD-4 and MD-5 versus MD-2 and MD-3. Notably, the highest estimated mean peak speed in training was approximately 93% of the estimated mean peak speed on match day for starters. Assuming that peak match speed is typically below the players' maximum speed, this finding suggests a potential shortfall in reaching maximum running speeds during training. In terms of load compensation for substitutes, differences between starters and substitutes in training duration, TD, peak speed, and $AccDec_{dist}$ on MD+2 were observed in most cycles, which could be due to residual fatigue from the last match in starters (Goulart et al., 2022) and/or recovery strategies from the coaching staff (Buchheit et al., 2021). However, there were no pronounced differences in HSRD and SpD, and overall load was considerably lower than on any other day also for non-starters.

Our results align with previous studies on female football players and male players, showing similar loading patterns and distances covered in relation to match days. However, there are limitations to our approach, including the lack of context surrounding each training day and the crude categorization of starters and non-starters.

5.4 Internal validity, external validity, and bias

Researchers needs to consider both internal and external validity when conducting or interpreting a study. Internal validity refers to the study's ability to measure what it wants to measure (Grimes & Schulz, 2002), which for this dissertation refers to the correctness of the estimates in each study. A second concern is the external validity, meaning whether the results can be extrapolated or generalized to real world settings (Grimes & Schulz, 2002).

The term bias in the context of research denotes deviation from the truth and undermines the internal validity of research (Grimes & Schulz, 2002). All observational studies have built in bias, thus a challenge for investigators is to ferret these out and judge how they might have affected results (Grimes & Schulz, 2002). One can broadly classify bias into two categories: selection bias and information bias.

5.4.1 Selection bias

Selection bias results from procedures used to select subjects and from factors that influence study participation (Rothman, 2012). Regarding the selection of our participants, it is quite evident that we did not perform a random selection from the base population of highly trained female football players, which would also be unrealistic. Rather, the teams were chosen based on the supervisors' pre-existing knowledge of the competence and consistency surrounding the sporting apparatus of Norwegian and Danish teams in close collaboration with Toppfotball Kvinner. This decision was rooted in the belief that prioritizing a more consistent data collection was preferable to random selection, wherein significant financial resources might be allocated to a team that, due to unforeseen factors, might not have contribute data as expected. In hindsight, the data collection period is a testament to the fact the future cannot be accurately predicted, as the Covid-19 epidemic was the root cause in destabilizing the infrastructure surrounding one of our teams, which in the end did not contribute data to a single study.

In addition to the selection process, there are several factors that might have influenced study participation. For example, we saw that changes in coaching staff resulted in less data for specific teams. The consistency in coaching staff could also be hypothesized to be dependent on success, suggesting that the study estimates are skewed towards teams that are successful, because these teams contributed to the bulk of non-missing data. Another contributing factor influencing study participation is the availability of club resources. For instance, one of the teams lacked a dedicated physical performance coach, significantly restricting the utilization of the tracking system and subsequently limiting the data collected. Consequently, these biases likely affect the studies' internal validity by skewing the estimates towards successful and resourceful teams characterized by consistent coaching staff that integrates tracking systems as an integral component of their training regimen. In addition, these biases also extend to external validity, limiting the broader applicability of the results to teams within this specific demographic.

5.4.2 Information bias

Grimes & Schultz (2002) states that information bias, also known as observation, classification, or measurement bias, results from incorrect determination of exposure or outcome, or both. External training load can be understood as exposure or dose in a causal framework (Impellizzeri et al., 2023). Meaning, a distal cause preceding a proximal cause (internal load), which in turn may result in beneficial training adaptations (Impellizzeri et al., 2023). Thus, it is important that the equipment used to collect external load (exposure) is both valid and reliable, and that any classification is based on physiological and performance considerations.

5.4.2.1 Validity of STATSports Apex 10 Hz

All studies utilized the STATSports Apex 10 Hz system, an athlete-tracking system released in August 2017 and widely adopted by professional clubs (Beato et al., 2018). This system uses 10 Hz multi-GNSS augmented units capable of acquiring and tracking multiple satellite systems (GPS, GLONASS, Galileo, and BeiDou) concurrently to provide accurate positional information. The validity of this system has only been investigated once in a peer-review study by Beato et al. (2018), and once by FIFA. In the study by Beato et al. (2018) the validity of the STATSports APEX system was tested against the criterion distance of 400 meters athletic track, a specific team sports circuit of 128 meters, and a 20-meter trial. In addition, the authors also assessed the validity of peak speed by comparing it against the criterion of a radar gun. The bias (%) in each trial was 1.05 ± 0.87%, 2.3 ± 1.1%, 1.11 ±

0.99%, and 2.36 ± 1.67%, respectively, which, according to Hopkins (2009) indicates a "good" rating. However, it should be noted that this study used a relatively simple field-based design, which is common in many validation studies of tracking systems (Luteberget & Gilgien, 2020). The issue with using distance as criteria is that they are not error free. As explained by Aughey (2011), the inherent error in the ability of a trundle wheel or tape measure to accurately measure distance, and difficulty in accurately determining the starting point for movement in the tracking system software. Another source of error is the difficulty in following the marked course for the participants (Linke et al., 2018). Hence, although the system's validity is supported by the study conducted by Beato et al. (2018), there remains a necessity for validation through more robust methodologies.

The authors also did not investigate the validity of acceleration metrics, which, as elaborated upon in the introduction section, is a key performance indicator in football. This has been examined most notably by Delaney et al. (2019) and Linke et al. (2018), albeit using a different brand of GPS (GPSports, Melbourne, Australia). However, this system also used the Doppler-method to calculate speed with a sampling rate of 10hz, meaning results can be somewhat extrapolated. Both studies also validated this system in sport-specific courses against a camera-based motion analysis system (VICON, Oxford, UK), which is a better criterion measure compared to set distance. Delaney et al. (2019) found that software-derived average acceleration showed larger bias than deriving this metric from the raw data, underlining the importance of considering smoothing/filtering techniques when processing data. Meanwhile, Linke et al. (2018) reported small to large differences in high accelerations ($\geq 3\ m \cdot s^{-2}$) and high decelerations ($\leq 3\ m \cdot s^{-2}$) during both a sport-specific course, a shuttle run, and a small-sided-game, when comparing said GPS system to VICON. Together, these studies add to the skepticism of acceleration and deceleration metrics by Buchheit et al. (2014), indicating that acceleration and deceleration metrics should be interpreted cautiously.

The STATSports APEX system is one of the few providers that has been approved by FIFAs Quality Programme for Electronic Performance Tracking Systems (FIFA, 2022). While not a peer-reviewed study, the testing protocols described in the report are quite rigorous and were performed by an independent academic institution (FIFA, 2022). The protocol included several test blocks: a circuit consisting of self-paced walking and jogging, maximal accelerations, and changes of direction; 2v2 and 5v5 small-sided games; sprints; and full pitch coverage. The protocol also used the motion capture system VICON, which is a better criterion compared to using set distances. The measure of accuracy was the root mean square difference (RMSD) between The STATSports APEX system and VICON; however, the exact value is never stated in the actual report. Instead, the system is rated on a z-scale based on the industry standard, where a score of 1.5 times the IQR is considered well-above industry standard. The STATSports APEX scored well-above industry standard in speed zones 0-7 and 7-15 $km \cdot h^{-1}$, and above industry standards for zones 15-20 and 20-25 $km \cdot h^{-1}$. However, the report states that data for >25$km \cdot h^{-1}$ are not available due to a lack of data collection, which is odd considering the inclusion of a sprint protocol.

In summary, the STATSports APEX system appears to be a valid tool for measuring distance and speed in football players, which strengthens the internal validity of our studies. However, special consideration needs to be taken when deriving acceleration metrics, as the validity of these metrics has not been rigorously tested and their accuracy is highly dependent on data processing.

5.4.2.2 GNSS signal quality and constellation

Bias can also result from factors affecting the quality of GNSS data. These factors have most notably been expanded upon by Malone et al. (2017) and includes signal quality, constellation of satellites, and sampling rate. Signal quality refers to the strength of the signal between the satellites and the device and may change depending on location and environmental obstruction (i.e., stadiums). Signal quality can be judged based on the number of satellites interacting with the receiver along with their orientation in the atmosphere, and should, always be recorded to ensure that longitudinal analysis can be carried out with confidence (Malone et al., 2017). GPS devices require a minimum of 4 satellites for adequate connection, with a higher number of connected satellites resulting in better coverage of the device (Larsson, 2003; J. J. Malone et al., 2017). Devices connecting to multiple GNSS, like STATSports APEX, have better coverage and signal strength compared to devices connecting to GPS only (Malone et al., 2017).

Another important factor is the horizontal dilution of precision (HDOP), which measures the accuracy of the horizontal positional signal determined by satellite geometry (Malone et al., 2017). As explained by Malone et al. (2017), when satellites are bunched together HDOP is high and precision is good, whereas when satellites are spread out HDOP is low, and precision is poor. Ideally, HDOP should be below 1, with values ranging from 0 to 50. Additionally, the sampling rate, i.e., the number of samples taken per second, plays an important role, with higher rates generally improving measurement precision.

In paper I and II, the mean ± SD number of satellites and HDOP was 17.5 ± 2.8 and 1.4 ± 0.6, respectively. In paper III, the mean HDOP ranged from 1 to 2, while the mean number of satellites ranged from 17.7 to 19.8. These findings, along with the sampling rate (10 Hz), suggests that the overall data quality in the project was sound, further bolstering the internal validity of our studies.

5.4.2.3 Data processing

Another source of bias that may have affected the results is the processing of the GNSS files. For example, many match files did not include precise indicators for when the match ended, or when a player was substituted. As such we had to approximate these end points using the available event data. We also chose the five outside positions (CD, FB, CM, WM, FW) based on convention. However, like Scott et al. (2020) we could have increased the granularity by splitting CM further into central defensive midfielder (CDM), CM, and central attacking midfielder (CAM). We also interpret playing position rather crudely, and it is highly likely that players within the same position can have different tactical roles. For example, one FB could have the tactical role of inverted wingback, while another could be a more complete wingback. It is highly likely that the estimates for the position FB does not capture the different activities of these roles, with the pure wingback covering more SpD compared to the inverted wingback or a pure FB (Ju et al., 2023). Consequently, this is an opportunity for future study.

5.4.2.4 Pre-selected (absolute) versus individual (relative) thresholds

The selected speed thresholds may have led to misclassification of activities, influencing the accuracy and interpretation of the data. For instance, a sprint threshold of 20 km·h^{-1} is not experienced the same by every player; for some, this may be closer to their maximal capacity than for others. In contrast, individualizing the thresholds based on a player's physiological ability, such as 80 or 90% of maximal sprinting speed or 100% of maximal aerobic speed, logically provides a more valid measure of effort. To illustrate, Abt & Lovell (2009) found that high-speed running thresholds based on the second ventilatory threshold were significantly lower than those used as the default setting within the ProZone match analysis system. This led the authors to conclude that distance run at high speeds can be substantially underestimated. In a subsequent study, Lovell & Abt (2013) identified a 41% difference in the high-speed distance covered between two players in the same position when using individualized zones based on ventilatory thresholds. In contrast, the absolute thresholds yielded negligible (5–7%) differences in total and high-speed distances covered. Together these findings highlight the importance of tailoring speed thresholds to individual physiological capacities for more accurate performance assessment.

However, there are several drawbacks to this approach. First, there is a lack of consensus concerning which tests and how many should be chosen to establish these thresholds (Bradley & Vescovi, 2015). Nevertheless, a study by Hunter et al. (2014) suggests avoiding the use of singular fitness characteristics to individualize thresholds and, instead, opting for a combination of players' anaerobic threshold, maximal aerobic speed, and maximal sprint speed characteristics. While this approach would be theoretically beneficial, the logistical challenges of testing multiple teams several times throughout the season would be considerable. Moreover, there is little evidence to suggest that adopting individualized speed thresholds adds significant value in determining the internal load of female football players (D. Scott & Lovell, 2018). For example, when testing the correlation between individualized thresholds using a range of methods, and RPE, Edwards training impulse score (TRIMP), and minutes spent above 80% of heart rate maximum, D. Scott & Lovell (2018) did not observe stronger correlations compared to those of absolute thresholds. Considering the logistical feasibility and the limited adoption of individualized threshold in high-level football (Akenhead & Nassis, 2016), it supports the use of absolute thresholds in our studies.

5.4.2.5 Sprinting

Our sprint threshold set at 20 km·h^{-1} was mostly inspired by a paper by Bradley & Vescovi (2015), wherein they consider both sex differences in activity profiles and normative test data as important indicators for where thresholds should be set. For example, in their paper they draw on comparisons of activity profiles of men's versus women's UEFA Champions League using the same thresholds (18-25 km·h^{-1} and > 25 km·h^{-1}) which shows that male players cover 986 and 200 meters, respectively, while female players cover 718 meters and 59 meters (Bradley et al., 2014). They also point to the fact that most sprints in football are initiated from a moving start and are only 5 meters in length (Di Salvo et al., 2009, 2010). Furthermore, they also highlight that the 70% percentile for a 15-meter flying sprint (5-meter run in, 15-meter sprint) is 25 km·h^{-1} (Vescovi, 2012), which obviously mean that a lot of the sprint efforts are not captured if this is the threshold. Instead, Bradley & Vescovi (2015) argue for a threshold set at 20 km·h^{-1} based on the minimum speed reached in a flying 5-meter sprint (5 meter run-in, 5 meter sprint) in the study by Vescovi, 2012. Limited corroborative data on sprint profiles in female soccer players have since been published. However, Haugen et al. (2014) compiled test records from the Norwegian Olympic Training Center of 165 female elite football players between 1995 and 2010. While the test records did not contain 5-meter split times, estimated 5-meter performance can still be generated by modelling sprint performance using the mono-exponential equation set forth by Furusawa (1927) and expanded upon by Jovanović and Vescovi (2022) and Jovanović (2023). When using the reported 10th percentile data in Haugen et al. (2014) the model estimates flying 5-meter sprint time to be 0.8 seconds (Figure 9B) resulting in an average speed over the segment of 22.5 km·h^{-1}. Considering that the speed difference between the minimum and the 10th percentile in Vescovi (2012) was 1 km·h^{-1} this gives a rough estimate of 21.5 km·h^{-1} as the minimal speed reached during a flying 5-meter sprint in Norwegian elite female football players. Furthermore, while the aim of Pedersen et al. (Pedersen et al., 2019) was to examine the effect of maximal strength training on sprint speed and jump height in high-level female football players (at level 2 and 3 in Norway), their data on 5-, 10-, and 15-meter split times is quite useful for modelling sprint performance. Using the mean pre-test performance of the intervention group, the model estimates a flying 5-meter performance of 0.83 seconds resulting in a mean speed of 21.7 km·h^{-1} over the segment (Figure 9C).

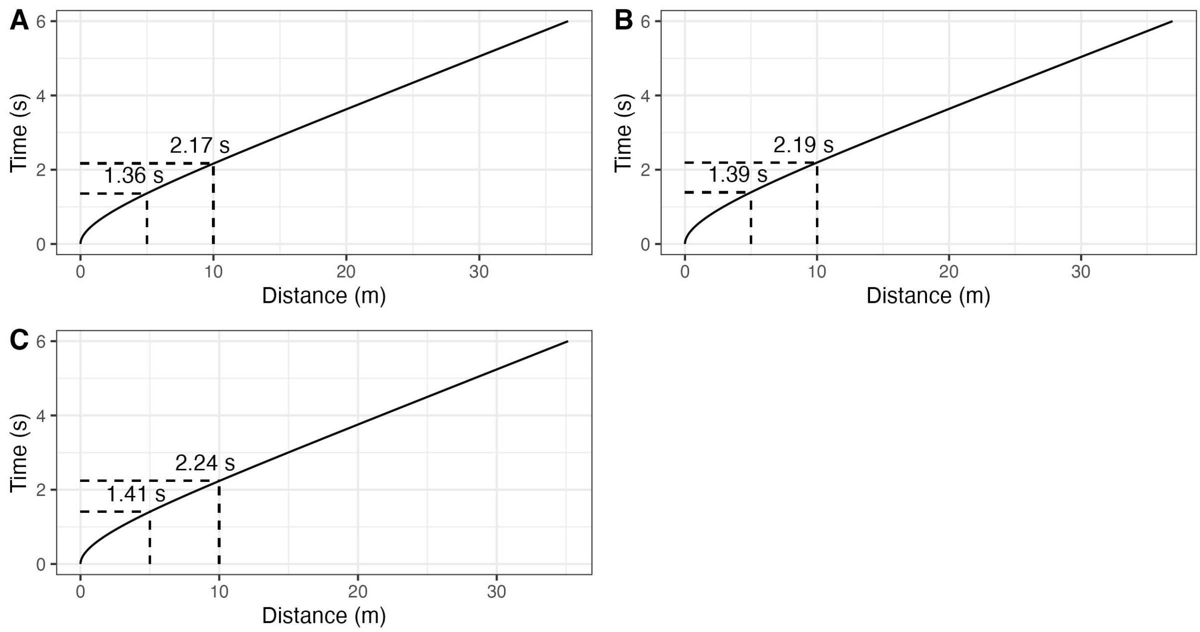

Figure 9: Modelled time~distance, corrected for starting procedure, using timing gate data from A) Vescovi (2012), B) Haugen et al. (2014), and C) Pedersen et al. (2019).

A point not considered by Bradley & Vescovi (2015) is the contrast in how speed is recorded in most GNSS-based tracking systems and how it is calculated when testing speed using timing gates. In most GNSS-based tracking systems Doppler-derived speed is recorded based on the sampling frequency, while speed from timing gates is calculated by dividing the known distance between two gates by the time it takes the athlete to traverse the distance. Thus, in the former approach speed can be recorded instantaneously, while the latter approach results in an average speed over the segment. The difference in the speed between these two approaches are only minor when measuring the maximal speed of an athlete, as this typically occurs between 20 and 30 meters for female football players where there is near zero acceleration (figure 10). In fact, both timing gates and 10 Hz GPS have showed to be valid tools for measuring maximal sprinting speed (Roe et al., 2017; Waldron et al., 2011). However, if we want to record a 5-meter flying sprint using a tracking system with the inherent assumption of a 5-meter lead-in from a static start, it might be more appropriate to set the threshold based on the instantaneous speed at 5 meters rather than the average speed between 5 and 10 meters. This would result in a slightly lower entry speed for sprinting (20.6 km·h^{-1}, 20.1 km·h^{-1}, 20.7 km·h^{-1}). Taken together, this approach supports the use of sprint thresholds set at 20 to 22.5 km·h^{-1}.

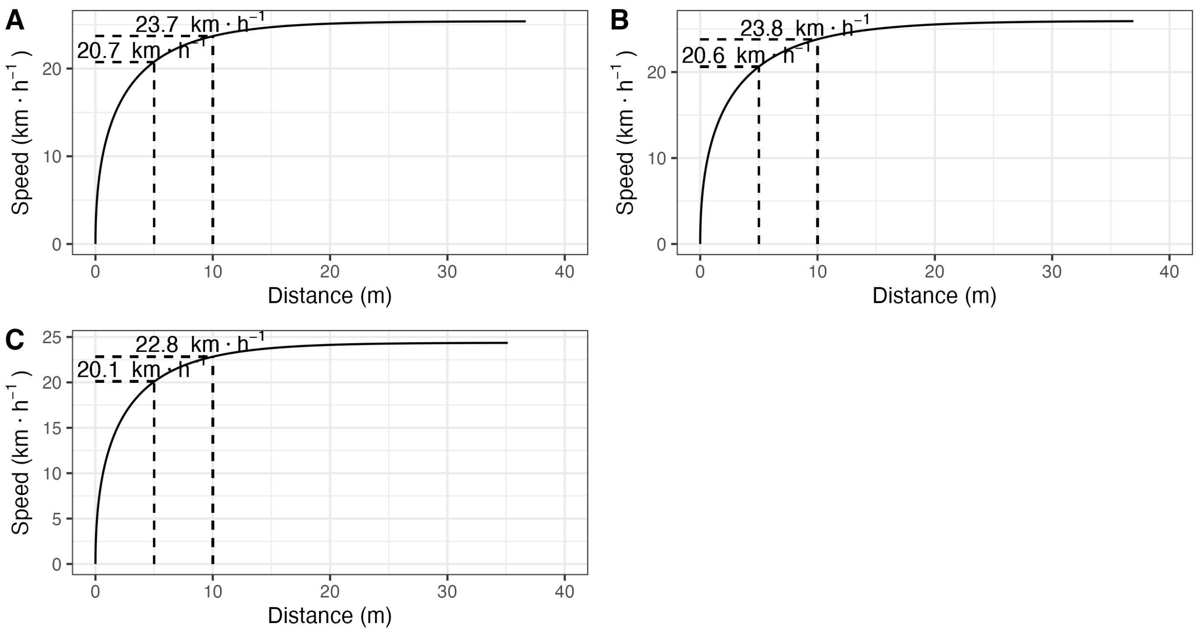

Figure 10: Modelled speed~distance, corrected for starting procedure, using data from A) Vescovi (2012), B) Haugen et al. (2014), and C) Pedersen et al. (2019).

5.4.2.6 High-speed running

Bradley & Vescovi (2015) also argue for the potential use of maximal aerobic speed when deciding upon individual or absolute thresholds for high-speed running. For instance, they point to evidence showing maximal aerobic speed of approximately 14.5 km·h^{-1} in elite female players (Haugen, Tønnessen, Hem, et al., 2014; Ingebrigtsen et al., 2011). The authors also highlight the use of intermittent field tests, specifically YoYo IR1, as having a potential for approximating maximal aerobic speed. As noted by the authors, these tests moderately to largely correlate with maximal aerobic speed (r = 0.69) and peak treadmill speed once VO_{2max} has been reached (r = 0.71) (Castagna et al., 2006). YoYo IR1 performance also significantly correlate (r = 0.762) with the total amount of high-intensity running (> 18 km·h^{-1}), and the amount of high-intensity running in the last 15 minutes of each half (r = 0.83) during match play in elite female players (Krustrup et al., 2005). The latter study also reported a mean YoYo IR1 value of 1379 meters, with a range from 600 to 1960 meters (Krustrup et al., 2015), which is equivalent to speeds of around 15.5 km·h^{-1} (range 14.5 – 16.5 km·h^{-1}) based on the test protocol in the original validity study by Krustrup et al. (2003). Similarly, Mujika et al.(2009) reported mean ± SD YoYo IR1 values of 1224 ± 255 in an elite cohort, equivalent to speeds of approximately 15.5 km·h^{-1}. Furthermore, based on the data from the English Football Association's national developmental program (479 elite youth and senior players), Datson et al. (2022) reported median YoYo IR1 of 1795 meters for 25 year old females, with 25th and 75th percentiles of 1480 and 2080 meters, corresponding to speed of 16, 15.5, and 16.5 km·h^{-1}, respectively. The lowest reported percentile in the dataset (0.9th percentile) for this group was 580 meters, equivalent to 14.5 km·h^{-1}.

Finally, to reiterate the reasoning of Bradley & Vescovi (2015), since intermittent bouts of high-intensity running during matches are associated with large anaerobic contributions and elevations in blood lactate concentration (Bangsbo et al., 2006), it seems logical that an absolute threshold should be situated around the onset of blood lactate accumulation. Thus, anchoring the threshold around the maximal aerobic speed, or YoYo test performance, appears to be a reasonable and sound approach, corresponding to 15-16 km·h^{-1} in elite female football players.

5.4.2.7 Acceleration and deceleration

We ended up using three different iterations of acceleration and deceleration metrics from the start to the end of the project. However, unlike 60% of studies (Ellens et al., 2022), and commensurate with best practice (Delves et al., 2021; J. J. Malone et al., 2017), we reported in each study the smoothing method, thresholds, and minimum effort duration used to ensure replicability.

In the first article we defined these metrics as *"...the distance covered with a positive or negative change in speed of more than ±2.26 m·s^{-2}, with a minimal effort duration of 0.3 s, finishing when the rate of acceleration/deceleration reached 0 m·s^{-2}"*. In contrast to most other studies, we chose to count the acceleration using three event markers. This allowed us to capture both the initial effort of accelerating or decelerating from 0 to ± 2.26 m·s^{-2}, and any subsequent submaximal acceleration or deceleration following exceeding the threshold. As explained by Varley et al. (2017), this may be a more practical definition for identifying acceleration deceleration efforts in contrast to purely quantifying the extremely short duration spent accelerating or deceleration above or below the required threshold and may better represent the perception of an acceleration held by practitioners. Furthermore, this method also allows the use of lower minimal effort durations (e.g. 0.3 seconds), as this ensures that a single effort will not be registered as multiple efforts (Varley et al., 2017). However, a possible drawback is that a new rise in acceleration or deceleration will not be registered as a new effort while the last event marker (0 m·s^{-2}) is not crossed (Varley et al., 2017).

The threshold of 2.26 m·s^{-2} was based on Trewin et al. (2018) wherein they state that this represents 80% of a player's acceleration over 10 meters during a 40-meter sprint testing. This was apparently determined during a pilot study. Unfortunately, this data was never published, and we had to take their word for it at the time. However, after checking against the modeled data from Haugen et al. (2014), Vescovi (2012), and Pedersen et al. (2019), this threshold does indeed seem plausible. In Table 6, one can see that the thresholds align with Pedersen et al. (2019), although it is somewhat lower than Haugen et al. (2014) and Vescovi (2012), which could be due to methodological dissimilarities.

It is also important to note that we derived these metrics from raw acceleration calculated over 0.6 seconds, which we did to provide a slight smoothing factor to the signal. The reason for choosing 0.6 seconds was that this resulted in each acceleration and deceleration lasting approximately 1.3 seconds, with players covering a mean distance of 6 meters. The reasoning here was that this matched what an acceleration would look like from a standing start, considering the pilot described by Trewin et al. (2018). In hindsight however, we should also have noted the initial- and end speed of each effort, as this would have provided more context to these metrics. The inherent assumptions are also based on accelerations from a standing start, abstracting away numerous accelerations while athletes are moving, and neglects to add any assumptions on decelerations.

Table 6: Time, time correction for starting procedure, speed, and 80% of acceleration at 10 meters.

Study	Time (s)	Time correction (s)	Speed ($m \cdot s^{-1}$)	80% of acceleration ($m \cdot s^{-2}$)
Haugen et al. 2014	2.19	0.40	6.61	2.42
Pedersen et al. 2019	2.24	0.35	6.34	2.27
Vescovi 2012	2.17	-0.12	6.58	2.43

For the second paper we switched to a) combining acceleration and deceleration into one metric, and b) reporting this as a count. The main reason for combining the two was for practical reasons. While accelerations and decelerations put distinct physiological and mechanical demands on the players (Harper et al., 2019), we considered it unlikely that these would be differentiated in a practical setting since so much of football training consists of specific exercises (Barrett et al., 2020). In addition, since there are so many ways these metrics can be reported (for the whole session, per minute, 1–10-minute peak periods, etc.) combining these metrics into one reduces the numbers of metrics reported by a factor of two. The reason for reporting this as a count was mostly to replicate the methodology by Oliva-Lozano et al. (2021) combined with the fact that most studies report accelerations and decelerations as a count (Figure 1E and 1F). Though, in hindsight we should have stuck with reporting this in meters, as counts treat each effort equally, irrespective of intensity or duration. To remedy this, we added combined acceleration and deceleration in meters to the results section of this dissertation.

In the final paper we defined combined acceleration and deceleration as *"…the distance covered with a positive or negative change in speed of more than ± 2.26 $m{\cdot}s^{-2}$, finishing when the rate of acceleration/deceleration reached 0 $m{\cdot}s^{-2}$"*. The decision to remove the minimal effort duration was based on a paper by Silva et al. (2023) showing that using minimal effort durations of 0.2 and 0.3 seconds underestimates accelerations and decelerations by 36% and 34%, respectively. Instead, the authors recommend that acceleration and deceleration efforts should be counted until the rate reaches 0 $m{\cdot}s^{-2}$, and any spikes in acceleration or deceleration passed the specified thresholds, however short, should be considered meaningful. We also switched back to expressing combined acceleration and deceleration in terms of distance instead of counts, simply because the former considers the intensity and duration of each effort. Moreover, acceleration was this time derived from speed smoothed using a 1 second rolling average. The decision to derive acceleration from smoothed speed was inspired by Elmer & Martin (2009) highlighting the importance of smoothing the original signal before differentiation, while the decision to use a rolling average over other smoothing methods (such as median or exponential filters) was due to its simplicity. The rolling window of 1 second was chosen after using match data as a pilot to model the mean distance, duration, initial speed, and final speed per acceleration and deceleration effort, along with the total count and distance at 94 minutes of playing time. As shown in Figure 11, this smoothing window results in estimates that beats the "eye" test. For example, using this smoothing window, most accelerations are initialized from walking speeds (4.1 km·h^{-1}), where the players accelerate a short distance (6.7 meters) in a short amount of time (2.2 seconds), before finishing when the speed crosses 15.2 km·h^{-1} .

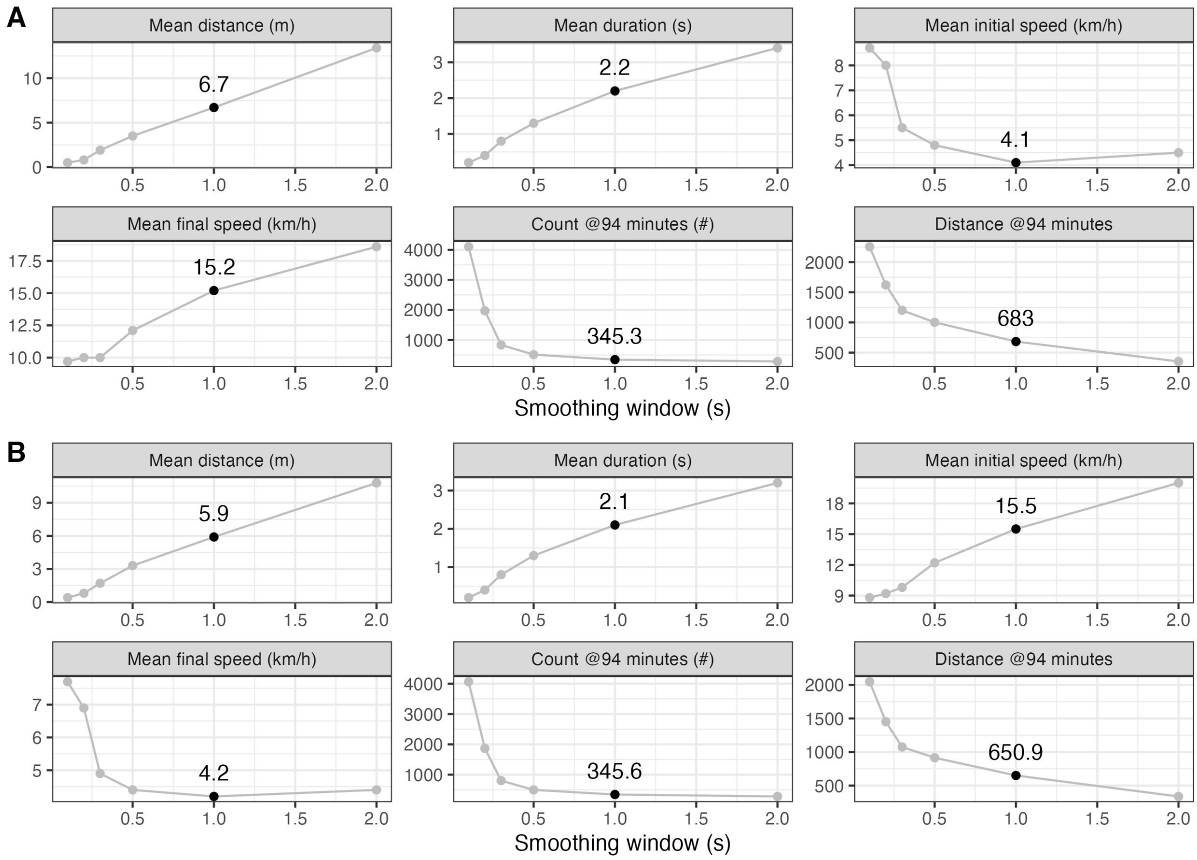

Figure 11: Effect of varying the rolling average window on A) acceleration, and B) deceleration metrics.

The utilization of three different iterations of acceleration and decelerations metrics throughout the project underscores a critical point: operationalizing these metrics into a seemingly valid framework proves to be exceptionally challenging, and there exists an extensively range of options in terms of how these metrics can be processed and defined. This challenge is particularly evident in Figure 11, which demonstrates how even slight variations in the smoothing window for a single method (rolling average) can severely impact the reported outputs. Therefore, it is imperative that future studies delve into the processing of acceleration and deceleration metrics to enhance their validity further.

5.4.2.8 Peak, next, and mean periods

We utilized epoch lengths of 1 and 5 minutes to describe peak, next, and mean periods. While these epoch lengths have consistently been used in the literature (Fransson et al., 2017; Harkness-Armstrong et al., 2022), the choice of exactly these two is rather arbitrary. For instance, in paper II we state that *"The epoch length for the peak locomotor demands was chosen according to the findings of Doncaster et al. (2020), where 1 min epochs produced the highest relative intensities when compared with 3- and 5-min epochs"*. However, this rational is not inherently sound, as distance or counts per minute decreases exponentially with increasing epoch length, as shown in Figure 9. Thus, using our own rationale, one might question why we did not use even shorter epoch length. Furthermore, one could argue that these epoch lengths are representative of the periods used by practitioners when prescribing work-to-rest ratios in small-sided games. However, using study protocols as a proxy for what practitioners are doing reveals that work periods of 4 minutes are mostly used (Hill-Haas et al., 2011).

A better approach would be to model the activity of peak periods as a function of epoch length using the formula: $Meters\ per\ minute = \beta_0 t^{-\beta_1}$, as proposed by Delaney et al. (2018) . For example, using the data in Riboli et al. (2021), a practitioner could calculate the 4-minute peak period by calculating $164 * 4^{-0.125}$, which equals 142 m·min^{-1} or 568 meters. Thus, future studies should instead adopt this approach when describing the peak periods of women's match play.

5.4.3 Generalizability (external validity)

The female players included in the studies were broadly classified as highly trained according to the definitions of McKay et al. (2022). Thus, our findings are more representative of this tier than others. In contrast to previous studies, our sample included data from four teams, which is a step forward with regards to reporting normative data for female football players, as most studies have utilized single-team designs (Harkness-Armstrong et al., 2022). The rating progression of the four teams in the final period of data collection is shown in Figure 12.

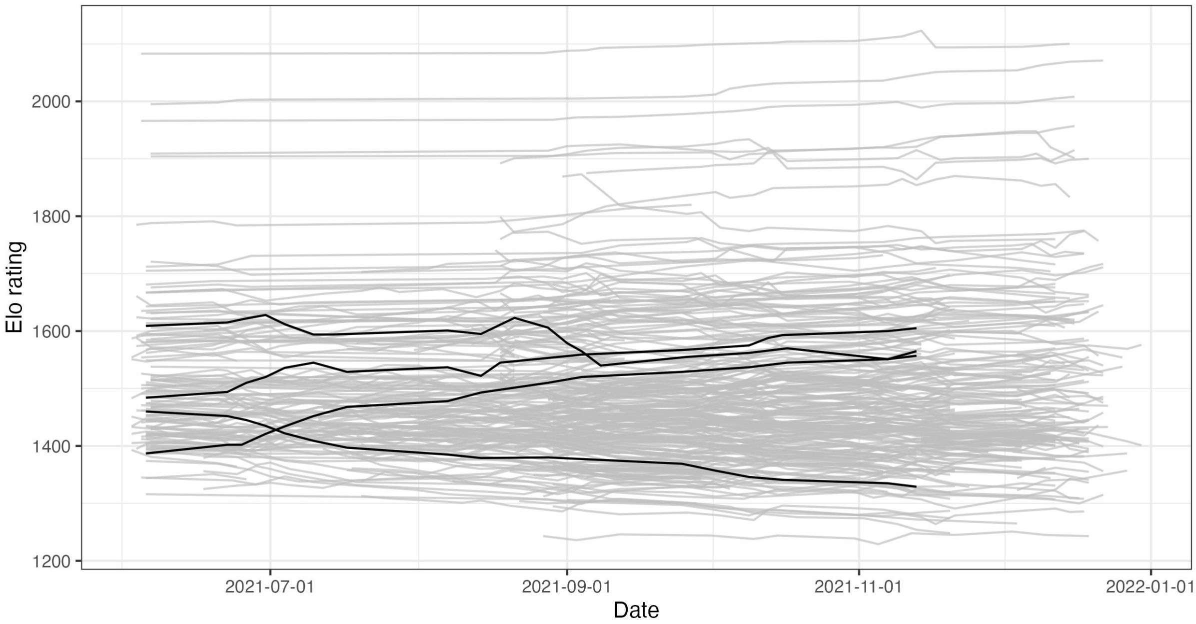

Figure 12: Elo rating for the four teams in the final period of data collection.

Furthermore, the data processing methods utilized in our research, including the selection of thresholds, epoch lengths, and smoothing techniques, heavily influence generalizability of the reported results. While we employed methods using a sound rational, variations in data processing approaches across studies may impact comparability. Practitioners should be aware that findings may not be directly applicable unless they use the exact same data processing methods as utilized in our study. Future research could benefit from standardizing data processing methods to enhance comparability and facilitate meta-analyses.

Additionally, the impact of playing styles and tactical approaches adopted by teams may influence the results observed in our study. League-specific variations in style of play and coaching may result in different activity profiles and training loads among players. Therefore, caution should be taken when generalizing our findings to teams employing different playing styles.

In summary, while our research provides valuable insights into the activity profiles and training loads of highly trained female football players, caution should be exercised when extrapolating these findings to broader populations. Future studies with larger and more diverse samples, standardized data processing methods, and consideration of playing styles and conditions are warranted to enhance the external validity and generalizability of findings in this field.

6 Conclusions, implications, and future research

6.1 Conclusions

Main conclusion 1: Activity profiles vary by playing position. 1 and 5-minute post-peak periods show minimal reductions in mean activity compared to mean 1 and 5-minute periods.

- The largest differences in activity are observed between CB and the wide positions (FB and WM), where the latter cover greater distances in terms of HSRD and SpD.
- Only trivial to small decreases in activity are observed between post-peak periods and corresponding mean match periods.

Main conclusion 2: There is considerable match-to-match variability in metrics that practitioners consider important.

- The observed match-to-match variability in HSRD, SpD, and $AccDec_{dist}$ ranges from 12 to 36%, while peak speed and TD ranges from 4.5 to 5%.

Main conclusion 3: Teams adjust their external load based on the number of days between matches, concentrating most of the load around mid-week in a pyramid-like fashion, while also minimally differentiating the load of starters and non-starters as match day approaches.

- The highest external loads occur three to four days before a match, succeeding and preceding days of lower load.
- Shorter cycles with fewer days between matches exhibited predominantly lower loads.
- Minimal differentiation in training load between starters and non-starters are observed from MD-5 onwards, irrespective of cycle length.
- Lower peak training speed relative to peak match speed, and a higher ratio of $AccDec_{dist}$ compared to SpD suggest preferences for small-sided games throughout the training week.

6.2 Practical implications

- The estimates provided in paper I can serve as normative data for practitioners, allowing them to compare the training and match activity of their own players against a large cohort of female football players. If used, practitioners are advised to consider the processing methods utilized for each metric, as numbers may not be universally applicable across all tracking systems. As a general recommendation, practitioners should consider both the mean values and confidence intervals when interpreting and scaling to our data.

- The considerable match-to-match variability in HSRD, SpD, and $AccDec_{dist}$ indicates that an extended monitoring period is needed to attain precision surrounding changes in these metrics. Conversely, the stability of TD and peak speed suggest that these metrics can reliably detect changes in match activity over a relatively shorter period.
- The pyramid-like distribution of the external training load aligns with established training principles, facilitating rest, loading, and peaking before a match. Coaches are advised to continue following this template.
- Practitioners should recognize the potential gap in achieving maximum running speed during training, as indicated by peak speeds reaching only approximately 93% of the estimated mean peak speed on match days. It's likely that this peak match speed is even lower than the players' maximal sprint speed. The relatively high ratio of $AccDec_{dist}$ to SpD further implies a preference for small-sided games, which alone may not provide a powerful enough stimulus for sprint adaptations. Practitioners should instead consider implementing sprint top-ups.

6.3 Future research

- We originally wanted to broaden the generalizability of the studies by including a Danish team in the cohort. However, due to unforeseen consequences of Covid-19, this team suffered an economic downturn with subsequent changes in coaching staff. Therefore, the data collection became insufficient, and we could not include them in the data material. However, incorporating and investigating the activity profiles of different leagues is an avenue for future study.
- There is a dire need for research attempting to validate and standardize processing methods for acceleration and deceleration metrics. The lack of consensus on processing methods hinders the comparability and generalizability of findings across studies.
- More studies should look at the interaction between determinants of performance, external load, and internal load, in accordance with the framework presented by Impellizzeri et al. (2019). Understanding these interactions can provide valuable insights into factors influencing player performance and injury risk, leading to more effective training and injury prevention strategies.

References

Abt, G., & Lovell, R. (2009). The use of individualized speed and intensity thresholds for determining the distance run at high-intensity in professional soccer. *Journal of Sports Sciences*, *27*(9), 893–898.

Ade, J., Drust, B., Morgan, O. J., & Bradley, P. S. (2021). Physiological characteristics and acute fatigue associated with position-specific speed endurance soccer drills: Production vs maintenance training. *Science and Medicine in Football*, *5*(1), 6–17.

Ade, J., Fitzpatrick, J., & Bradley, P. S. (2016). High-intensity efforts in elite soccer matches and associated movement patterns, technical skills and tactical actions. Information for position-specific training drills. *Journal of sports sciences*, *34*(24), 2205–2214.

Akenhead, R., & Nassis, G. P. (2016). Training load and player monitoring in high-level football: Current practice and perceptions. *International Journal of Sports Physiology and Performance*, *11*(5), 587–593.

Allen, T., Taberner, M., Zhilkin, M., & Rhodes, D. (2023). Running more than before? The evolution of running load demands in the English Premier League. *International Journal of Sports Science & Coaching*, *0*(0). https://doi.org/10.1177/17479541231164507

Andersson, H. Å., Randers, M. B., Heiner-Møller, A., Krustrup, P., & Mohr, M. (2010). Elite female soccer players perform more high-intensity running when playing in international games compared with domestic league games. *The Journal of Strength & Conditioning Research*, *24*(4), 912–919.

API Football. (2023). *API-FOOTBALL (3.9.2)*. https://www.api-football.com/documentation-v3

Aughey, R. J. (2011). Applications of GPS technologies to field sports. *International Journal of Sports Physiology and Performance*, *6*(3), 295–310.

Bache-Mathiesen, L. K., Andersen, T. E., Clarsen, B., & Fagerland, M. W. (2022). Handling and reporting missing data in training load and injury risk research. *Science and Medicine in Football*, *6*(4), 452–464.

Bangsbo, J., Mohr, M., & Krustrup, P. (2006). Physical and metabolic demands of training and match-play in the elite football player. *Journal of Sports Sciences*, *24*(07), 665–674.

Baptista, I., Johansen, D., Figueiredo, P., Rebelo, A., & Pettersen, S. A. (2020). Positional Differences in Peak-and Accumulated-Training Load Relative to Match Load in Elite Football. *Sports*, *8*(1), 1.

Barrett, S., Varley, M. C., Hills, S. P., Russell, M., Reeves, M., Hearn, A., & Towlson, C. (2020). Understanding the influence of the head coach on soccer training drills—An 8 season analysis. *Applied Sciences*, *10*(22), 8149.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beato, M., Coratella, G., Stiff, A., & Iacono, A. D. (2018). The validity and between-unit variability of GNSS units (STATSports Apex 10 and 18 Hz) for measuring distance and peak speed in team sports. *Frontiers in Physiology*, *9*, 1288.

Borg, D. N., Nguyen, R., & Tierney, N. J. (2022). Missing data: Current practice in football research and recommendations for improvement. *Science and Medicine in Football*, *6*(2), 262–267.

Bradley, P. S., Dellal, A., Mohr, M., Castellano, J., & Wilkie, A. (2014). Gender differences in match performance characteristics of soccer players competing in the UEFA Champions League. *Human Movement Science*, *33*, 159–171.

Bradley, P. S., & Noakes, T. D. (2013). Match running performance fluctuations in elite soccer: Indicative of fatigue, pacing or situational influences? *Journal of Sports Sciences*, *31*(15), 1627–1638.

Bradley, P. S., & Vescovi, J. D. (2015). Velocity thresholds for women's soccer matches: Sex specificity dictates high-speed-running and sprinting thresholds—Female Athletes in Motion (FAiM). *International Journal of Sports Physiology and Performance*, *10*(1), 112–116.

Buchheit, M., Al Haddad, H., Simpson, B. M., Palazzi, D., Bourdon, P. C., Di Salvo, V., & Mendez-Villanueva, A. (2014). Monitoring accelerations with GPS in football: Time to slow down? *International Journal of Sports Physiology and Performance*, *9*(3), 442–445.

Buchheit, M., Lacome, M., Cholley, Y., & Simpson, B. M. (2018). Neuromuscular responses to conditioned soccer sessions assessed via GPS-embedded accelerometers: Insights into tactical periodization. *International Journal of Sports Physiology and Performance*, *13*(5), 577–583.

Buchheit, M., Sandua, M., Berndsen, J., Shelton, A., Smith, S., Norman, D., McHugh, D., & Hader, K. (2021). Loading patterns and programming practices in elite football: Insights from 100 elite practitioners. *Sport Performance & Science Reports*, *153*, 1–18.

Buchheit, M., Settembre, M., Hader, K., & McHugh, D. (2023). Exposures to near-to-maximal speed running bouts during different turnarounds in elite football: Association with match hamstring injuries. *Biology of Sport*, *40*(4), 1057–1067.

Buchheit, M., Settembre, M., Hader, K., & McHugh, D. (2024). From High-Speed Running to

Hobbling on Crutches: A Machine Learning Perspective on the Relationships Between Training Doses and Match Injury Trends. *Sport Performance & Science Reports*, *216*, 1–11.

Buchheit, M., & Simpson, B. M. (2017). Player-tracking technology: Half-full or half-empty glass? *International Journal of Sports Physiology and Performance*, *12*(s2), S2-35-S2-41.

Cardinale, M., & Varley, M. C. (2017). Wearable training-monitoring technology: Applications, challenges, and opportunities. *International Journal of Sports Physiology and Performance*, *12*(s2), S2-55-S2-62.

Castagna, C., Impellizzeri, F. M., Chamari, K., Carlomagno, D., & Rampinini, E. (2006). Aerobic fitness and yo-yo continuous and intermittent tests performances in soccer players: Acorrelation study. *The Journal of Strength & Conditioning Research*, *20*(2), 320–325.

Choice, E., Tufano, J., Jagger, K., Hooker, K., & Cochrane-Snyman, K. C. (2022). Differences across Playing Levels for Match-Play Physical Demands in Women's Professional and Collegiate Soccer: A Narrative Review. *Sports*, *10*(10), 141.

Clemente, F. M., Martins, F. M., & Mendes, R. S. (2014). Periodization based on small-sided soccer games: Theoretical considerations. *Strength & Conditioning Journal*, *36*(5), 34–43.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Academic press.

Coutts, A. J., Crowcroft, S., & Kempton, T. (2018). *Developing athlete monitoring systems: Theoretical basis and practical applications*.

Dalen, T., Jørgen, I., Gertjan, E., Havard, H. G., & Ulrik, W. (2016). Player load, acceleration, and deceleration during forty-five competitive matches of elite soccer. *The Journal of Strength & Conditioning Research*, *30*(2), 351–359.

Datson, N., Weston, M., Drust, B., Atkinson, G., Lolli, L., & Gregson, W. (2022). Reference values for performance test outcomes relevant to English female soccer players. *Science and Medicine in Football*, *6*(5), 589–596.

de Araújo, M. C., & Mießen, K. A. (2017). Twenty years of the FIFA women's world cup: An outstanding evolution of competitiveness. *Women in Sport and Physical Activity Journal*, *25*(1), 60–64.

De Hoyo, M., Cohen, D. D., Sañudo, B., Carrasco, L., Álvarez-Mesa, A., Del Ojo, J. J., Domínguez-Cobo, S., Mañas, V., & Otero-Esquina, C. (2016). Influence of football match time–motion parameters on recovery time course of muscle damage and jump ability. *Journal of Sports Sciences*, *34*(14), 1363–1370.

Delaney, J. A., Thornton, H. R., Rowell, A. E., Dascombe, B. J., Aughey, R. J., & Duthie, G. M. (2018). Modelling the decrement in running intensity within professional soccer players.

Science and Medicine in Football, *2*(2), 86–92.

Delaney, J. A., Wileman, T. M., Perry, N. J., Thornton, H. R., Moresi, M. P., & Duthie, G. M. (2019). The validity of a global navigation satellite system for quantifying small-area team-sport movements. *The Journal of Strength & Conditioning Research*, *33*(6), 1463–1466.

Delves, R. I., Aughey, R. J., Ball, K., & Duthie, G. M. (2021). The quantification of acceleration events in elite team sport: A systematic review. *Sports Medicine - Open*, *7*(1), 1–35.

Di Salvo, V., Baron, R., González-Haro, C., Gormasz, C., Pigozzi, F., & Bachl, N. (2010). Sprinting analysis of elite soccer players during European Champions League and UEFA Cup matches. *Journal of sports sciences*, *28*(14), 1489–1494.

Di Salvo, V., Gregson, W., Atkinson, G., Tordoff, P., & Drust, B. (2009). Analysis of high intensity activity in Premier League soccer. *International Journal of Sports Medicine*, 205–212.

Diaz-Seradilla, E., Rodríguez-Fernández, A., Rodríguez-Marroyo, J. A., Castillo, D., Raya-González, J., & Villa Vicente, J. G. (2022). Inter-and intra-microcycle external load analysis in female professional soccer players: A playing position approach. *PLoS ONE*, *17*(3), e0264908. https://doi.org/10.1371/journal.pone.0264908

Draganidis, D., Chatzinikolaou, A., Avloniti, A., Barbero-Álvarez, J. C., Mohr, M., Malliou, P., Gourgoulis, V., Deli, C. K., Douroudos, I. I., & Margonis, K. (2015). Recovery kinetics of knee flexor and extensor strength after a football match. *PLoS ONE*, *10*(6), e0128072.

Edouard, P., Mendiguchia, J., Guex, K., Lahti, J., Samozino, P., & Morin, J.-B. (2019). Sprinting: A potential vaccine for hamstring injury. *Sport Performance & Science Reports*, *48*, 1–2.

Ekstrand, J., Hallén, A., Marin, V., & Gauffin, H. (2023). Most modifiable risk factors for hamstring muscle injury in women's elite football are extrinsic and associated with the club, the team, and the coaching staff and not the players themselves: The UEFA Women's Elite Club Injury Study. *Knee Surgery, Sports Traumatology, Arthroscopy*, *31*, 2550–2555.

Ellens, S., Middleton, K., Gastin, P. B., & Varley, M. C. (2022). Techniques to derive and clean acceleration and deceleration data of athlete tracking technologies in team sports: A scoping review. *Journal of Sports Sciences*, *40*(16), 1772–1800.

Elmer, S. J., & Martin, J. C. (2009). Fourier Series Approximations and Low Pass Filtering: Facilitating Learning of Digital Signal Processing for Biomechanics Students. *Sportscience*, *13*, 1–8.

Faude, O., Koch, T., & Meyer, T. (2012). Straight sprinting is the most frequent action in goal situations in professional football. *Journal of Sports Sciences*, *30*(7), 625–631.

Fereday, K., Hills, S. P., Russell, M., Smith, J., Cunningham, D. J., Shearer, D., McNarry, M., & Kilduff, L. P. (2020). A comparison of rolling averages versus discrete time epochs for assessing the worst-case scenario locomotor demands of professional soccer match-play. *Journal of Science and Medicine in Sport*, *23*(8), 764–769.

FIFA. (2022). *FIFA EPTS Test Report STATSports*.

Fransson, D., Krustrup, P., & Mohr, M. (2017). Running intensity fluctuations indicate temporary performance decrement in top-class football. *Science and Medicine in Football*, *1*(1), 10–17.

Furusawa, K., Hill, A. V., & Parkinson, J. L. (1927). The dynamics of" sprint" running. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character*, *102*(713), 29–42.

Fyfe, J. J., Bishop, D. J., & Stepto, N. K. (2014). Interference between concurrent resistance and endurance exercise: Molecular bases and the role of individual training variables. *Sports Medicine*, *44*, 743–762.

Gastin, P. B., Hunkin, S. L., Fahrner, B., & Robertson, S. (2019). Deceleration, acceleration, and impacts are strong contributors to muscle damage in professional Australian football. *The Journal of Strength & Conditioning Research*, *33*(12), 3374–3383.

Gaudino, P., Alberti, G., & Iaia, F. M. (2014). Estimated metabolic and mechanical demands during different small-sided games in elite soccer players. *Human Movement Science*, *36*, 123–133.

Goulart, K. N. O., Coimbra, C. C., Campos, H. O., Drummond, L. R., Ogando, P. H. M., Brown, G., Couto, B. P., Duffield, R., & Wanner, S. P. (2022). Fatigue and recovery time course after female soccer matches: A systematic review and meta-analysis. *Sports Medicine-Open*, *8*(1), 1–21.

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: A guide to misinterpretations. *European Journal of Epidemiology*, *31*, 337–350.

Grimes, D. A., & Schulz, K. F. (2002). Bias and causal associations in observational research. *The Lancet*, *359*(9302), 248–252.

Hader, K., Mendez-Villanueva, A., Palazzi, D., Ahmaidi, S., & Buchheit, M. (2016). Metabolic power requirement of change of direction speed in young soccer players: Not all is what it seems. *PLoS ONE*, *11*(3), e0149839.

Harkness-Armstrong, A., Till, K., Datson, N., Myhill, N., & Emmonds, S. (2022). A systematic review of match-play characteristics in women's soccer. *PLoS ONE*, *17*(6), e0268334.

Harper, D. J., Carling, C., & Kiely, J. (2019). High-intensity acceleration and deceleration demands in elite team sports competitive match play: A systematic review and meta-analysis of observational studies. *Sports Medicine*, *49*, 1923–1947.

Harper, D. J., & Kiely, J. (2018). Damaging nature of decelerations: Do we adequately prepare players? *BMJ Open Sport & Exercise Medicine*, *4*(1), e000379.

Haugen, T. A., Breitschädel, F., & Seiler, S. (2020). Sprint mechanical properties in soccer players according to playing standard, position, age and sex. *Journal of Sports Sciences*, *38*(9), 1070–1076.

Haugen, T. A., Tønnessen, E., Hem, E., Leirstein, S., & Seiler, S. (2014). VO2max characteristics of elite female soccer players, 1989–2007. *International Journal of Sports Physiology and Performance*, *9*(3), 515–521.

Haugen, T. A., Tønnessen, E., Hisdal, J., & Seiler, S. (2014). The role and development of sprinting speed in soccer. *International Journal of Sports Physiology and Performance*, *9*(3), 432–441.

Hill-Haas, S. V., Dawson, B., Impellizzeri, F. M., & Coutts, A. J. (2011). Physiology of small-sided games training in football: A systematic review. *Sports Medicine*, *41*, 199–220.

Hopkins, W. G. (1997). *A new view of statistics*. Will G. Hopkins.

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine*, *30*, 1–15.

Hopkins, W. G. (2004). How to interpret changes in an athletic performance test. *Sportscience*, *8*, 1–7.

Hopkins, W. G. (2017). Spreadsheets for analysis of validity and reliability. *Sportscience*, *21*.

Hopkins, W. G., Hawley, J. A., & Burke, L. M. (1999). Design and analysis of research on sport performance enhancement. *Medicine & Science in Sports & Exercise*, *31*(3), 472–485.

Hopkins, W. G., Marshall, S., Batterham, A., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Medicine & Science in Sports & Exercise*, *41*(1), 3.

Hunter, F., Bray, J., Towlson, C., Smith, M., Barrett, S., Madden, J., Abt, G., & Lovell, R. (2014). Individualisation of time-motion analysis: A method comparison and case report series. *International Journal of Sports Medicine*, 41–48.

Hvattum, L. M., & Arntzen, H. (2010). Using ELO ratings for match result prediction in association football. *International Journal of forecasting*, *26*(3), 460–470.

Iaia, F. M., & Bangsbo, J. (2010). Speed endurance training is a powerful stimulus for physiological adaptations and performance improvements of athletes. *Scandinavian journal of medicine & science in sports*, *20*, 11–23.

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.

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.

Ingebrigtsen, J., Dillern, T., & Shalfawi, S. A. (2011). Aerobic capacities and anthropometric characteristics of elite female soccer players. *The Journal of Strength & Conditioning Research*, *25*(12), 3352–3357.

Jovanović, M. (2023). Bias in estimated short sprint profiles using timing gates due to the flying start: Simulation study and proposed solutions. *Computer Methods in Biomechanics and Biomedical Engineering*, 1–11.

Jovanović, M., & Vescovi, J. (2022). {shorts}: An R Package for Modeling Short Sprints. *International Journal of Strength and Conditioning*, *2*(1).

Ju, W., Doran, D., Hawkins, R., Evans, M., Laws, A., & Bradley, P. (2023). Contextualised high-intensity running profiles of elite football players with reference to general and specialised tactical roles. *Biology of Sport*, *40*(1), 291–301.

Karlsson, U. B., Vagle, M., Wiig, H., & Luteberget, L. S. (2023). Training Load Quantification in Women's Elite Football: A Season-Long Prospective Cohort Study. *International Journal of Sports Physiology and Performance*, *18*(6), 603–614.

Kirkendall, D. T., & Krustrup, P. (2022). Studying professional and recreational female footballers: A bibliometric exercise. *Scandinavian journal of medicine & science in sports*, *32*, 12–26.

Krustrup, P., Mohr, M., Amstrup, T., Rysgaard, T., Johansen, J., Steensberg, A., Pedersen, P. K., & Bangsbo, J. (2003). The yo-yo intermittent recovery test: Physiological response, reliability, and validity. *Medicine & Science in Sports & Exercise*, *35*(4), 697–705.

Krustrup, P., Mohr, M., Ellingsgaard, H., & Bangsbo, J. (2005). Physical demands during an elite female soccer game: Importance of training status. *Medicine & Science in Sports & Exercise*, *37*(7), 1242–1248.

Krustrup, P., Mohr, M., Nybo, L., Draganidis, D., Randers, M. B., Ermidis, G., Ørntoft, C., Røddik, L., Batsilas, D., & Poulios, A. (2022). Muscle metabolism and impaired sprint performance in an elite women's football game. *Scandinavian Journal of Medicine & Science in Sports*, *32*, 27–38.

Lacome, M., Simpson, B. M., Cholley, Y., Lambert, P., & Buchheit, M. (2018). Small-sided games in elite soccer: Does one size fit all? *International Journal of Sports Physiology and Performance*, *13*(5), 568–576.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863.

Larsson, P. (2003). Global positioning system and sport-specific testing. *Sports Medicine*, *33*, 1093–1101.

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). *Emmeans: Estimated marginal means, aka least-squares means. R package version 1.5*.

Linke, D., Link, D., & Lames, M. (2018). Validation of electronic performance and tracking systems EPTS under field conditions. *PLoS ONE*, *13*(7), e0199519.

Lovell, R., & Abt, G. (2013). Individualization of time–motion analysis: A case-cohort example. *International Journal of Sports Physiology and Performance*, *8*(4), 456–458.

Luteberget, L. S., & Gilgien, M. (2020). Validation methods for global and local positioning-based athlete monitoring systems in team sports: A scoping review. *BMJ Open Sport & Exercise Medicine*, *6*(1), e000794.

Magnusson, A., Skaug, H., Nielsen, A., Berg, C., Kristensen, K., Maechler, M., van Bentham, K., Bolker, B., Brooks, M., & Brooks, M. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, *9*(2), 378–400.

Malcata, R. M. (2014). *Modelling Progression of Competitive Sport Performance*. Auckland University of Technology.

Malcata, R. M., & Hopkins, W. G. (2014). Variability of competitive performance of elite athletes: A systematic review. *Sports Medicine*, *44*, 1763–1774.

Malone, J. J., Lovell, R., Varley, M. C., & Coutts, A. J. (2017). Unpacking the black box: Applications and considerations for using GPS devices in sport. *International Journal of Sports Physiology and Performance*, *12*(S2), S218–S226.

Malone, S., Owen, A., Mendes, B., Hughes, B., Collins, K., & Gabbett, T. J. (2018). High-speed running and sprinting as an injury risk factor in soccer: Can well-developed physical qualities reduce the risk? *Journal of Science and Medicine in Sport*, *21*(3), 257–262.

McCall, A., Pruna, R., Van der Horst, N., Dupont, G., Buchheit, M., Coutts, A. J., Impellizzeri, F. M., & Fanchini, M. (2020). Exercise-Based strategies to prevent muscle injury in male elite footballers: An Expert-Led Delphi survey of 21 practitioners belonging to 18 teams from the Big-5 European Leagues. *Sports Medicine*, *50*, 1667–1681.

McKay, A. K., Stellingwerff, T., Smith, E. S., Martin, D. T., Mujika, I., Goosey-Tolfrey, V. L., Sheppard, J., & Burke, L. M. (2022). Defining Training and Performance Caliber: A Participant Classification Framework. *International Journal of Sports Physiology and Performance*, *17*(2), 317–331.

Mendez-Villanueva, A., Buchheit, M., Simpson, B. E. N., Peltola, E. S. A., & Bourdon, P. (2011). Does on-field sprinting performance in young soccer players depend on how fast they can run or how fast they do run? *The Journal of Strength & Conditioning Research*, *25*(9), 2634–2638.

Mohr, M., Krustrup, P., Andersson, H., Kirkendal, D., & Bangsbo, J. (2008). Match activities of elite women soccer players at different performance levels. *The Journal of Strength & Conditioning Research*, *22*(2), 341–349.

Morgans, R., Orme, P., Anderson, L., & Drust, B. (2014). Principles and practices of training for soccer. *Journal of Sport and Health Science*, *3*(4), 251–257.

Mujika, I., Santisteban, J., Impellizzeri, F. M., & Castagna, C. (2009). Fitness determinants of success in men's and women's football. *Journal of Sports Sciences*, *27*(2), 107–114.

Norsk Teknisk Byrå (NTB). (2023). *Norsk Internasjonal Fotballstatistikk*. nifs.no

Nosek, P., Brownlee, T. E., Drust, B., & Andrew, M. (2021). Feedback of GPS training data within professional English soccer: A comparison of decision making and perceptions between coaches, players and performance staff. *Science and Medicine in Football*, *5*(1), 35–47.

Novak, A. R., Impellizzeri, F. M., Trivedi, A., Coutts, A. J., & McCall, A. (2021). Analysis of the worst-case scenarios in an elite football team: Towards a better understanding and application. *Journal of Sports Sciences*, *39*(16), 1850–1859.

Oliva-Lozano, J. M., Muyor, J. M., Fortes, V., & McLaren, S. J. (2021). Decomposing the variability of match physical performance in professional soccer: Implications for monitoring individuals. *European Journal of Sport Science*, *21*(11), 1588–1596.

Pedersen, S., Heitmann, K. A., Sagelv, E. H., Johansen, D., & Pettersen, S. A. (2019). Improved maximal strength is not associated with improvements in sprint time or jump height in high-level female football players: A cluster-randomized controlled trial. *BMC Sports*

Science, Medicine and Rehabilitation, *11*(1), 20.

Rampinini, E., Alberti, G., Fiorenza, M., Riggio, M., Sassi, R., Borges, T. O., & Coutts, A. J. (2014). Accuracy of GPS devices for measuring high-intensity running in field-based team sports. *International Journal of Sports Medicine*, 49–53.

Riboli, A., Coratella, G., Rampichini, S., Cé, E., & Esposito, F. (2020). Area per player in small-sided games to replicate the external load and estimated physiological match demands in elite soccer players. *PLoS ONE*, *15*(9), e0229194.

Riboli, A., Esposito, F., & Coratella, G. (2022). The distribution of match activities relative to the maximal intensities in elite soccer players: Implications for practice. *Research in Sports Medicine*, *30*(5), 463–474.

Riboli, A., Francini, L., Rossi, E., Caronti, A., Boldrini, L., & Mazzoni, S. (2023). Top-class women's soccer performance: Peak demands and distribution of the match activities relative to maximal intensities during official matches. *Biology of Sport*, *41*(1), 207–215.

Roe, G., Darrall-Jones, J., Black, C., Shaw, W., Till, K., & Jones, B. (2017). Validity of 10-HZ GPS and timing gates for assessing maximum velocity in professional rugby union players. *International Journal of Sports Physiology and Performance*, *12*(6), 836–839.

Romero-Moraleda, B., Nedergaard, N. J., Morencos, E., Casamichana, D., Ramirez-Campillo, R., & Vanrenterghem, J. (2021). External and internal loads during the competitive season in professional female soccer players according to their playing position: Differences between training and competition. *Research in Sports Medicine*, *29*(5), 449–461.

Rothman, K. J. (2012). *Epidemiology: An introduction*. Oxford University Press.

Rothman, K. J., & Greenland, S. (2018). Planning study size based on precision rather than power. *Epidemiology*, *29*(5), 599–603.

Scott, D., Haigh, J., & Lovell, R. (2020). Physical characteristics and match performances in women's international versus domestic-level football players: A 2-year, league-wide study. *Science and Medicine in Football*, *4*(3), 211–215.

Scott, D., & Lovell, R. (2018). Individualisation of speed thresholds does not enhance the dose-response determination in football training. *Journal of Sports Sciences*, *36*(13), 1523–1532.

Scott, M. TU., Scott, T. J., & Kelly, V. G. (2016). The validity and reliability of global positioning systems in team sport: A brief review. *The Journal of Strength & Conditioning Research*, *30*(5), 1470–1490.

Silva, H., Nakamura, F. Y., Ribeiro, J., Asian-Clemente, J., Roriz, P., & Marcelino, R. (2023).

Using minimum effort duration can compromise the analysis of acceleration and deceleration demands in football. *International Journal of Performance Analysis in Sport*, 1–13.

Stevens, T. G., de Ruiter, C. J., Twisk, J. W., Savelsbergh, G. J., & Beek, P. J. (2017). Quantification of in-season training load relative to match load in professional Dutch Eredivisie football players. *Science and Medicine in Football*, *1*(2), 117–125.

Torres-Ronda, L., Beanland, E., Whitehead, S., Sweeting, A., & Clubb, J. (2022). Tracking systems in team sports: A narrative review of applications of the data and sport specific analysis. *Sports Medicine - Open*, *8*(1), 1–22.

Trewin, J., Meylan, C., Varley, M. C., & Cronin, J. (2018). The match-to-match variation of match-running in elite female soccer. *Journal of Science and Medicine in Sport*, *21*(2), 196–201.

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*, 1–67.

Vanrenterghem, J., Nedergaard, N. J., Robinson, M. A., & Drust, B. (2017). Training load monitoring in team sports: A novel framework separating physiological and biomechanical load-adaptation pathways. *Sports Medicine*, *47*, 2135–2142.

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 and Performance*, *12*(8), 1059–1068.

Vescovi, J. D. (2012). Sprint speed characteristics of high-level American female soccer players: Female Athletes in Motion (FAiM) study. *Journal of Science and Medicine in Sport*, *15*(5), 474–478.

Viru, A., & Viru, M. (2000). Nature of training effects. *Exercise and Sport Science*, *6795*.

Waldron, M., Worsfold, P., Twist, C., & Lamb, K. (2011). Concurrent validity and test–retest reliability of a global positioning system (GPS) and timing gates to assess sprint performance variables. *Journal of Sports Sciences*, *29*(15), 1613–1619.

Weaving, D., Young, D., Riboli, A., Jones, B., & Coratella, G. (2022). The maximal intensity period: Rationalising its use in team sports practice. *Sports Medicine - Open*, *8*(1), 1–9.

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength & Conditioning Research*, *19*(1), 231–240.

Paper I

Received: 23 October 2021 | Revised: 9 November 2021 | Accepted: 17 November 2021

DOI: 10.1111/sms.14105

ORIGINAL ARTICLE

WILEY

Position specific physical performance and running intensity fluctuations in elite women's football

A. K. Winther[1] | I. Baptista[2,3] | S. Pedersen[1] | M. B. Randers[1,4] | D. Johansen[3] | P. Krustrup[4,5,6,7] | S. A. Pettersen[1]

[1]School of Sport Sciences, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway

[2]Faculty of Sport, University of Porto, Porto, Portugal

[3]Department of Computer Science, Faculty of Science and Technology, UiT The Arctic University of Norway, Tromsø, Norway

[4]Department of Sports Science and Clinical Biomechanics, SDU Sport and Health Sciences Cluster (SHSC), University of Southern Denmark, Odense, Denmark

[5]Shangai University of Sport, Shanghai, China

[6]Sport and Health Sciences, University of Exeter, Exeter, UK

[7]Danish Institute for Advanced Study (DIAS), University of Southern Denmark, Odense, Denmark

Correspondence
A. K. Winther, School of Sport Sciences, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway.
Email: andreas.k.winther@uit.no

Funding information
UiT - The Arctic University of Norway; Tromsø Research Foundation

Abstract

The purpose of the present study was to investigate the physical performance of elite female football players during match play along with transient alterations in running performance following 1- and 5-min univariate peak periods. 54 elite female players from four top-level Norwegian teams were monitored for one season ($n = 393$ match observations), and physical performance data collected using STATSport GPS APEX. Results revealed significant differences in physical performance between the positions during full match play, particularly between wide and central players. Both full backs (FBs) and wide midfielders (WMs) covered more total distance (TD), high-speed running distance (HSRD), and sprint distance (SpD) than center backs (CBs) ($p < 0.05$–0.001), while WMs also covered more HSRD than both central midfielders (CMs) ($p < 0.01$) and forwards (FWs) ($p < 0.05$), and more acceleration -and deceleration distance (Acc_{dist} and Dec_{dist}) than both CBs and CMs ($p < 0.01$–0.001). A similar pattern was observed for the peak period analysis, with FBs and WMs covering more SpD in peak 1 min than CBs and CM ($p < 0.001$) and more SpD in peak 5-min than CBs, CMs, and FWs ($p < 0.001$). Irrespective of the variable analyzed, greater distances were covered during the peak 5-min period than in the next-5 and mean 5-min periods ($p < 0.001$). Significant ($p < 0.001$), but small to trivial (Cohen's D_z: 0.07–0.20), decreases in distance covered were also observed for each variable following each univariate peak 5-min period. In conclusion, practitioners should account for differences in physical performance when developing training programs for female football players and be aware of transient reductions in physical performance following univariate peak 1- and 5-min periods. Specifically, the very high intensity in 1-min peak periods adds support to the principal of executing speed endurance activities during training to mirror and be prepared for the physical demands of match play.

KEYWORDS

global positioning system, peak periods, physical performance, women's football

1 | INTRODUCTION

Women's football has surpassed an undeniable transformation during the last decade, and its development has been a priority for the Fédération Internationale de Football Association.[1] This increased professionalism and growing popularity have impacted the scientific community with focused research increasing the body of knowledge regarding the women's game. Nevertheless, studies about player positioning monitoring and match physical performance are still scarce, since most of the research topics in women's football are related to injury.[2]

Time-motion analysis involving the intermittent activity pattern of women's football is necessary to assess the locomotor and mechanical demands of match play, which in turn is essential for specific training prescription.[3,4] Women's football has been described as a sport with multiple brief intense actions separated by low-intensity activities, with mean values for total distance (TD) and high-speed running distance (HSRD) ranging from 9.2–11.3 km to 1.2–2.7 km, respectively.[5–7] However, it is well documented in male football that different playing positions accumulate different external match load[8–11] and that such load presents large individual variations.[4,12] Therefore, to describe and characterize physical demands of football competitions, it is recommended to present these analyses by playing positions rather than reporting only the team averages.[13]

The majority of the studies that aim to analyze the external load of match play through locomotor activity do not account the energy cost associated with accelerations (Acc) and decelerations (Dec),[14] which may underestimate match load by 6%–8%.[15,16] To the best of our knowledge, only three studies women's football[17–19] have included the metrics of Acc and Dec in their analysis, while simultaneously adopted a more detailed categorization of the playing positions (into 4–6 positions) instead of the commonly used categorization into defenders, midfielders, and attackers.[4,5,7,20–23] However, the study of Mara et al.[17] included a considerably small sample size (12 players across 7 matches) and their intention was to focus only on Acc and Dec profiles, excluding other important variables such as HSRD and sprints from the analysis.

The reporting of absolute or average demands has been advantageous to profile the players' overall physical loading. However, it must be noted that football presents a stochastic nature[24] and training programs designed to replicate these average demands of competition will likely lead to players being underprepared for the more intense periods of a football match.[25] While high-intensity phases have received particular attention in men's football in recent years,[26–32] sparse information has been provided in relation to the peak demands of different playing positions in women's football. Another interesting aspect is whether decrements in high-intensity running occur following these periods, which may be indicative of physiological fatigue or pacing strategy.[33] However, while several studies on men have found transient decrements following high-intensity phases of 1 and 5 min,[34,35] no study to date has investigated this in women.

The most intense periods have been studied using different methodologies, including different temporal durations (epochs) and analysis techniques. Studies initially started by examining fixed-time periods of 15[24,36] or 5 min.[33,35] However, in a systematic review of the methodologies used to quantify the peak match demands, Whitehead et al.[37] concluded that pre-defined time periods lack sensitivity to find the true peaks of physical outputs when compared with a rolling average method. Indeed, in a study with elite male football players, Varley et al.[38] reported that fixed compared with rolling 5-min epochs underestimated peak running demands by up to 25%, which is in line with more recent research that also analyzed shorter time periods (eg, 1 and 3 min).[27,35,39] Despite Trewin et al.[19] having studied the most intense periods in match play of elite female football players using a rolling average approach, the authors only analyzed 5-min epochs, resulting in limited information for training prescription.[37]

Therefore, the aims of the present study were twofold. We first aimed to characterize the physical performance in elite women's football by position. Secondly, we aimed to investigate transient alterations in running demands following rolling peak periods of 1 and 5 min.

2 | METHODS

With ethical institutional approval from the Norwegian Centre for Research Data (reference number: 296155) and written informed consent from the participants, 108 female football players (22 ± 4 years of age) from four top-level Norwegian clubs were included in the study. Locomotor data from the four clubs′ official matches in the 2020 season (60 matches) were collected using GPS APEX (STATSports), with a sampling frequency of 10 Hz. The validity and levels of accuracy (bias <5%) of this tracking system have been previously presented.[40] During matches, each player wore a tight vest with the GPS unit on the back of their upper body between scapula as described by the manufacturer. The microsensor devices were activated 15 min prior to the start of each match, in accordance with the manufacturer's recommendations and previous research,[41] with this period of time being excluded from analyses. To minimize inter-devices error,[40] each player used the same GPS unit during the entire season.

Doppler derived speed data was exported from manufacturer software (STATSport Sonra 2.1.4) into Python 3.7.6. for processing (linearly interpolating any missing raw data) and to derive metrics. Raw acceleration was then calculated over a period of 0.6 s. After deriving all the metrics, the data were transferred to R (R.4.0.5, R Core Team, 2021) for statistical analysis.

2.1 | Physical performance variables

The physical parameters analyzed included total distance (TD), high-speed running distance (HSRD) (>4.44 m.s^{-1}), sprint distance (SpD) (>5.55 m.s^{-1}), acceleration and deceleration distances (Acc_{dist}/Dec_{dist}), and peak speed ($Peak_{speed}$). Acc_{dist} and Dec_{dist} were defined as the distance covered with a positive or negative change in speed of more than $\pm$2.26 m.s^{-2}, with a minimal effort duration of 0.3 s, finishing when the rate of acceleration/deceleration reached 0 m.s^{-2}. The speed thresholds were chosen according to the previous research.[19,20] Except for $Peak_{speed}$, all other variables were used to analyze both full match (absolute values) and peak locomotor demands (1- and 5-min peak periods rolling analysis periods). The epoch length for the peak locomotor demands was chosen according to the findings of Doncaster et al.,[39] where 1-min epochs produced the highest relative intensities when compared with 3- and 5-min epochs.

2.2 | Statistical analysis

Both between-positional differences during full match and within-positional differences between peak, next, and mean periods, were determined using linear mixed-modelling. To deal with the nested structure of the data, we treated matches in which two of our teams met as separate matches, and, due to positional differences in locomotor demands, the same player in a new position as a new player. Furthermore, to get a representative sample, we only included players who completed, at least, two full-time (90 min) matches. Also, match performance data of <90 min were treated as missing, and goalkeepers were excluded from analysis. This resulted in an initial sample of 501 observations with 108 missing values, which were subsequently removed in the complete case analysis (CCA). The final sample included 393 match observations (M_{obs}) from 54 players (center backs, CB, $n = 10$, $M_{obs} = 113$; full backs, FB, $n = 11$, $M_{obs} = 84$; central midfielders, CM, $n = 16$, $M_{obs} = 105$; wide midfielders, WM, $n = 9$; $M_{obs} = 57$ and forwards, FW, $n = 8$, $M_{obs} = 34$). These positions were chosen according to previous research.[35] The mean number of satellites and horizontal dilution of precision was 17.5 $\pm$ 2.8 and 1.4 $\pm$ 0.6, respectively. For the full match between-positional analysis, we specified for each physical parameter a model with *Position* as the fixed effect and *Team*, *Match ID*, and *Player ID*, as the random effects. Similarly, to investigate within-positional differences between peak, next, and mean periods, we specified for each physical parameter a model with *Position*, *Period*, and an interaction term as the fixed effects, and *Team*, *Match ID*, and *Player ID*, as the random effects. Moreover, the Tukey method was applied to adjust the multiple comparisons, with an α-level set at 0.05 as the level of significance. To calculate effect sizes (ES) we used Cohen's D_z.[42] All statistical analyses were done using the lme4[43] and emmeans[44] packages. Unless otherwise stated all results are estimate marginal means $\pm$ 90% confidence intervals.

3 | RESULTS

3.1 | Full match activity profiles

There were significant differences between certain playing positions across all metrics except for peak speed (Table 1). The results obtained for TD and HSRD revealed that CB covered less distance than both FB and CM. Moreover, also WM performed higher HSRD than CM and FW. Regarding sprint distance, CB covered less distance than FB, WM, and CM, with WM also presenting higher values than FW. Significant higher values were also observed for WM than FW for total distance and high-speed distance (Table 1).

No significant differences in peak speed were observed between outfield positions. Regarding the acceleration profiles, WM performed higher Acc_{dist} than CB and CM, and higher Dec_{dist} than both CB, CM, and FW (Table 1).

3.2 | 1- and 5-min peak period profiles

No significant differences were observed between positions in 1-min peaks for TD. However, three playing positions (FB, WM, and CM) performed significantly higher peak 5-min TD compared with CB (Table 2). FB and WM performed more 1- and 5-min peak HSRD than CB during both periods, with WM also performing more HSRD than CM and FW in the 5-min peak (Table 2). The results obtained for SpD revealed a similar trend, with FB and WM presenting higher values in the 1-min peak, than CB and CM, and in the 5-min peak than CB, CM, and FW. WM was the playing position with the highest values observed

TABLE 1 Full match activity profiles by position

	CB	FB	CM	WM	FW	Contrasts
TD (m)	8934 ± 264	9590 ± 255	9982 ± 229	10131 ± 284	9376 ± 311	FB > CB (656 ± 557)*; WM > CB (1197 ± 591)*; WM > FW (755 ± 646)*; CM > CB (1048 ± 525)*
HSRD (m)	1054 ± 148	1573 ± 144	1483 ± 130	1894 ± 160	1429 ± 174	FB > CB (519 ± 308)***; WM > CB (840 ± 327)***; WM > CM (411 ± 300)**; WM > FW (465 ± 359)*; CM > CB (429 ± 290)**
SpD (m)	227 ± 54	413 ± 53	293 ± 47	530 ± 59	380 ± 65	FB > CB (187 ± 118)*; WM > CB (303 ± 126)***; WM > CM (237 ± 116)***; FW > CB (154 ± 134)*
Peak speed (km/h)	27 ± 1	28 ± 1	27 ± 1	29 ± 1	28 ± 1	No sig. differences
Acc (m)	427 ± 42	488 ± 41	433 ± 36	578 ± 46	506 ± 51	WM > CB (151 ± 97)**; WM > CM (145 ± 90)**
Dec (m)	305 ± 34	406 ± 33	361 ± 30	493 ± 38	382 ± 42	FB > CB (101 ± 75)**; WM > CB (188 ± 80)***; WM > CM (132 ± 73)***; WM > FW (111 ± 88)*

*$p < 0.05$.; **$p < 0.01$.; ***$p < 0.001$.

for Acc_{dist} and Dec_{dist}, both in 1- and 5-min peak periods, with results being significantly higher, during 1-min peak, than CM, and higher than CB, CM, and FW during 5-min peak (Table 2).

3.3 | Running intensity fluctuations (peak, next and mean periods)

Both CB, FB, CM, and FW presented significantly higher values during the 1-min peak than in the following 5-min periods, for HSRD, SpD, $Acc_{dist,}$ and Dec_{dist} (Figure 1). A similar trend was seen for WM, who also presented significantly higher peak 1-min versus next 5-min values, except for HSRD. Furthermore, small but significant decreases in distance covered. Furthermore, both CB, FB, and WM covered less distance, during the 5-min period following the peak 1-min compared to the mean 5-min period. For CM, there were no differences between these two epochs in Acc_{dist}, while for FW the same was observed in TD, HSRD, and SpD. With exception of TD, CB presented significantly higher values during the peak 1-min period compared to the mean 5-min period. Similarly, FB and CM presented higher SpD and Acc_{dist} during the 1-min peak. For WM and FW, significantly differences between those moments were observed only in SpD.

With respect to the analysis of peak, next, and mean 5-min, the same trend, without exception, was observed for every playing position (Figure 2). Irrespectively of the variable analyzed, the results revealed higher intensities during the peak 5-min than in both next and mean 5-min periods. Next 5-min periods also presented lower values compared to the mean 5-min of each variable (ES range: 0.07–0.20).

4 | DISCUSSION

For the first time, running intensity fluctuations using 1- and 5-min peak periods have been studied in detail in elite women's football. The major findings are that that HSRD, Acc_{dist} and Dec_{dist} in the 1-min peak correspond to ~50% of the distances covered in the 5-min peak and that the peak 1-min sprint period is significantly higher, in every playing position, than the mean 5-min period for the same variable. In addition, these differences between 1- and 5-min peaks are even smaller in SpD, with the most demanding minute of the match corresponding to ≥60% of the SpD performed in the 5-min peak.

These findings are in line with previous research in professional male footballers[29] and may be important for practitioners during training prescription. As an example, it may allow coaches to make evidence-based decisions regarding durations for exercises that aim to replicate, or to prepare, the players to cope with these peak periods of the match. Preparing players to cope with the 5-min peak periods of the match do not necessarily mean that these players will be ready for the most demanding 1-min peaks,

TABLE 2 Peak period (1 and 5 min) profiles by position

	CB	FB	CM	WM	FW	Contrasts
Peak 1-min period						
TD (m)	174 ± 15	192 ± 16	189 ± 14	191 ± 19	178 ± 23	No sig. differences
HSRD (m)	71 ± 9	93 ± 9	85 ± 9	93 ± 11	77 ± 12	FB > CB (22 ± 15)**; WM > CB (21 ± 16)*
SpD (m)	37 ± 4	53 ± 4	40 ± 4	54 ± 5	44 ± 6	FB > CB (16 ± 7)***; FB > CM (13 ± 7)***; WM > CB (18 ± 8)***; WM > CM (14 ± 8)***
Acc (m)	28 ± 2	32 ± 3	28 ± 2	34 ± 3	31 ± 3	WM > CM (6 ± 5)*
Dec (m)	20 ± 2	24 ± 2	21 ± 2	27 ± 2	23 ± 3	WM > CB (7 ± 4)*; WM > CM (6 ± 4)**
Peak 5-min period						
TD (m)	634 ± 21	688 ± 22	706 ± 20	712 ± 26	658 ± 31	FB > CB (54 ± 37)**; WM > CB (78 ± 41)***; CM > CB (72 ± 35)***
HSRD (m)	139 ± 13	190 ± 13	179 ± 12	210 ± 14	164 ± 16	FB > CB (52 ± 21)***; WM > CB (71 ± 23)***; WM > CM (30 ± 21)**; WM > FW (45 ± 26)***; CM > CB (41 ± 20)***
SpD (m)	54 ± 6	82 ± 6	63 ± 6	92 ± 7	67 ± 8	FB > CB (28 ± 11)***; FB > CM (19 ± 10)***; FB > FW (15 ± 12)*; WM > CB (38 ± 11)***; WM > CM (29 ± 11)***; WM > FW (25 ± 13)***
Acc (m)	56 ± 4	66 ± 4	56 ± 3	74 ± 4	62 ± 5	FB > CB (10 ± 7)**; FB > CM (10 ± 6)**; WM > CB (18 ± 7)***; WM > CM (17 ± 7)***; WM > FW (12 ± 8)**
Dec (m)	41 ± 3	50 ± 3	45 ± 3	59 ± 3	46 ± 4	FB > CB (10 ± 5)***; FB > CM (6 ± 5)*; WM > FB (9 ± 5)***; WM > CB (19 ± 5)***; WM > CM (14 ± 5)***; WM > FW (13 ± 6)***

$^{*}p < 0.05$.; $^{**}p < 0.01$.; $^{***}p < 0.001$.

since the demands of 5-min peaks are not evenly distributed across every minute.

Interestingly, the performance in the 5-min period following the peak 5-min in SpD is similar to the performance observed after the peak 1-min, suggesting that the 1-min peak period is so physically demanding that it requires a long recovery period with lower intensity. Furthermore, the high intensity in the SpD 1-min peak period adds support to the prescription of speed endurance activities during training to mirror and be prepared for the physical demands of match play.[45,46]

Corroborating previous studies regarding the presence and development of temporary fatigue[47] after peak periods,[29,48] our results revealed a significant decrease of high-intensity actions in the 5-min period following the peak 1-min, across several playing positions. The next 5-min period was also less demanding, in every variable (except for Acc_{dist}), than the 5-min rolling average, for CB, FB, CM, and WM. However, while this decrease was significant, it is important to note that the differences in distance covered were quite small and that post 5-min periods are quite variable.[19]

It is important to have reference values by playing position for the demands of match play in elite women's football, since comparisons to men's football are not commensurable. To date, only two other studies[18,19] have simultaneously described the distribution of both running and acceleration patterns in elite women's football. In our study, apart from TD, a pattern emerged in the full match analysis in which external positions covered more distance in all speed zones, compared with central positions. This was especially apparent for SpD where both FB and WM covered significantly more distances than CB and CM, which partly supports the conclusions of Panduro et al.[18] where CB was considered the playing position with the lowest overall physical match demands. A similar trend was observed in the analysis of the 5-min peak periods, where FB and WM presented the highest values in high-speed variables, while CB was the playing position with the lowest work-rate in every variable analyzed. These results are somewhat similar with previous research in elite male[8] and female[13] footballers; however, in the study of Panduro et al.,[18] the authors reported CM as one of the most demanding playing positions regarding high-speed activities,

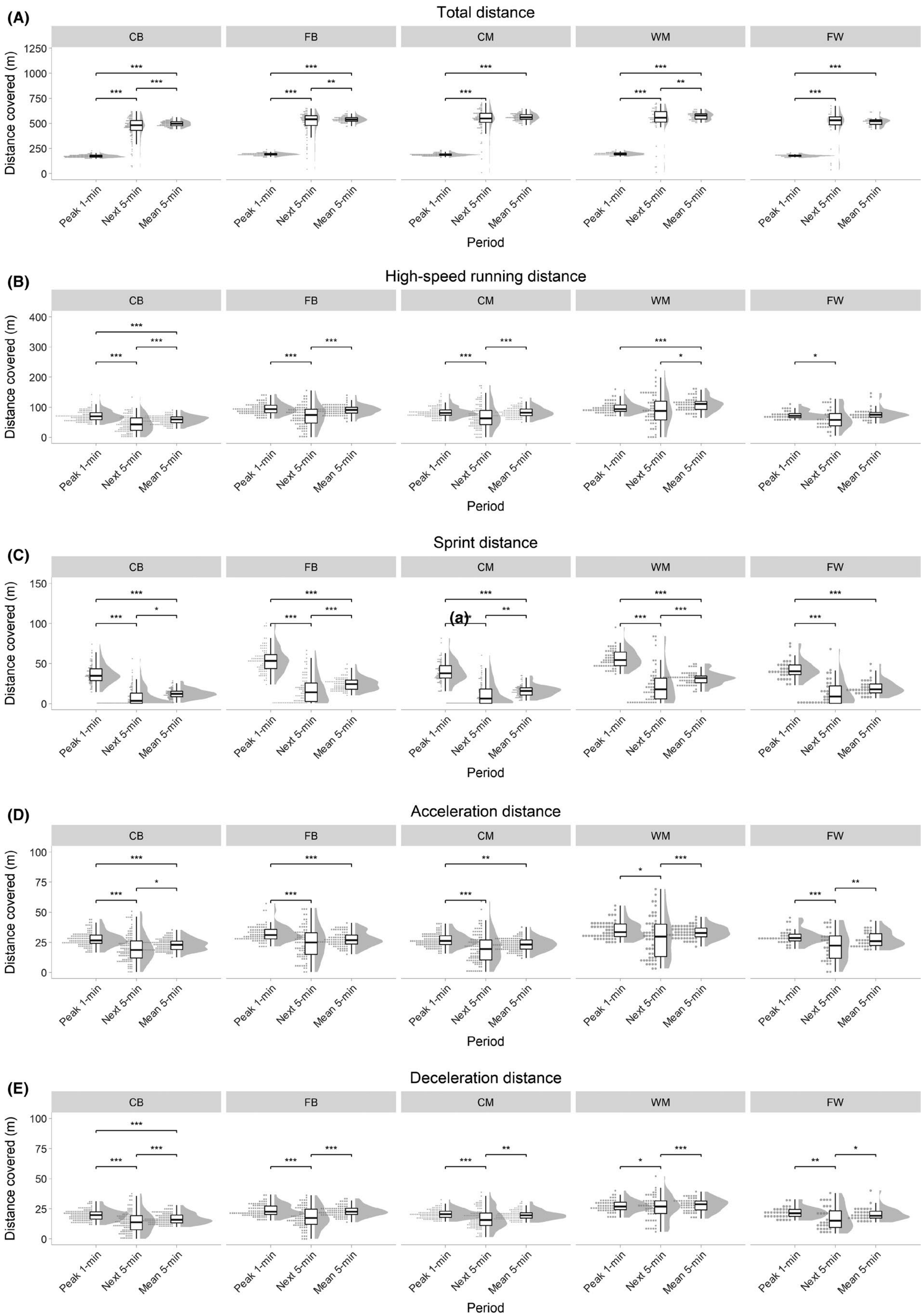

FIGURE 1 Distance covered during the peak 1-min, the next 5-min, and the mean 5-min period, for total distance (A), high-speed distance (B), sprint distance (C), acceleration distance (D), and deceleration distance (E)

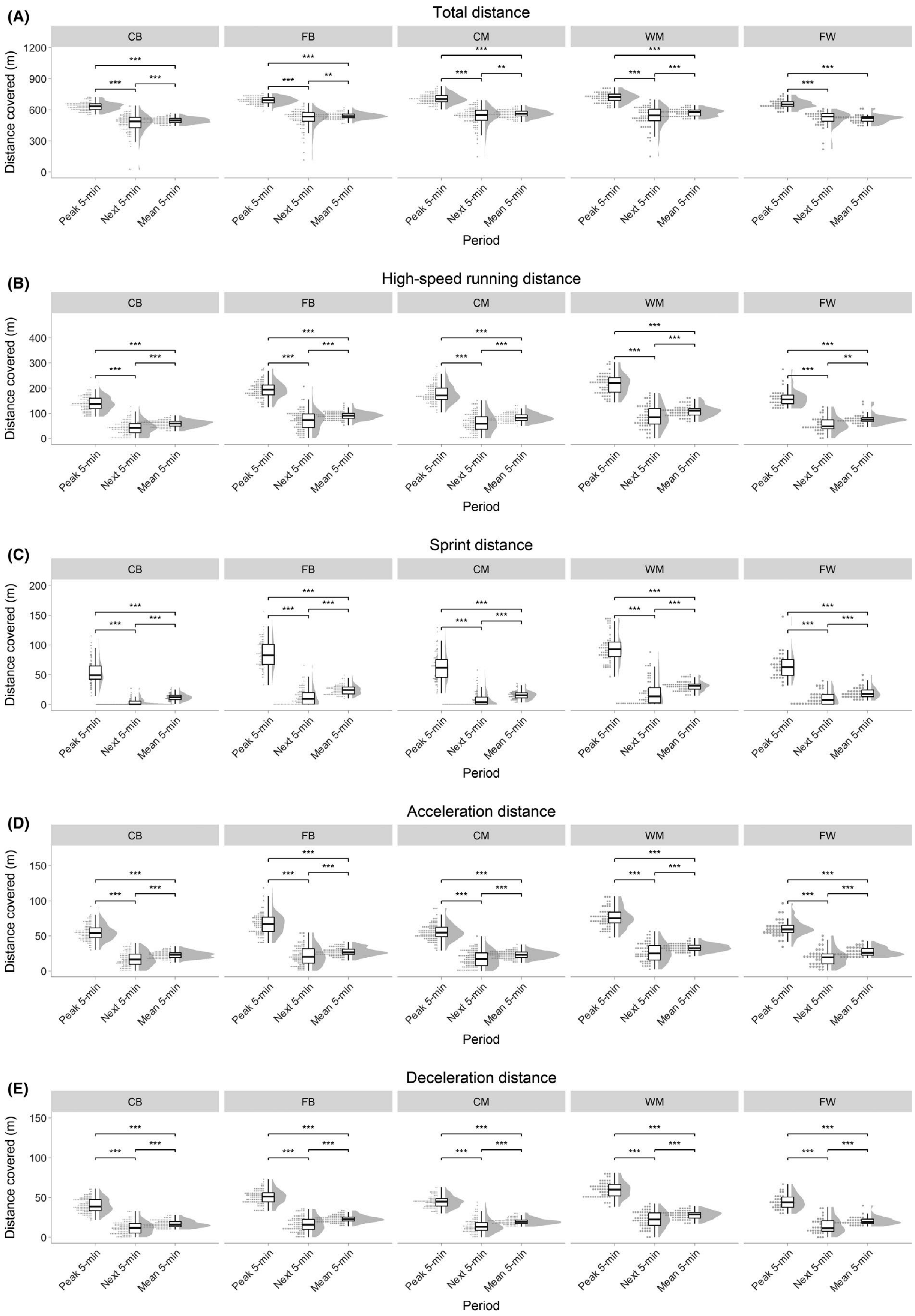

FIGURE 2 Distance covered during the peak 5-min, the next 5-min, and the mean 5-min period, for total distance (A), high-speed distance (B), sprint distance (C), acceleration distance (D), and deceleration distance (E)

which is not in line with the findings of this research. In fact, 5-min peaks present larger differences between positions than 1-min peaks, which may be explained by the accumulation of differences within 5 min. The three studies used different tracking systems, and direct comparisons between studies should be done with care.

This study gathered performance data from top quality players (three teams ranked Top-4 in the National League), resulting in a large dataset, which is both rare and novel in studies on elite athletes. However, the dataset was not evenly distributed across playing positions, with FW presenting a considerably smaller sample size than the other positions. In fact, the inclusion criteria chosen for the present study (players had to play the full match—90 min) together with the fact that FW were the players more often substituted in match, resulted in a smaller sample size for this group and hence lower statistical power for the running intensity fluctuation analysis.

5 | PERSPECTIVES

The results of this study emphasize that peak 1-min SpD in all positions and Acc- and Dec distance in some positions are significantly higher than the mean 5-min period in these variables, which should have implications in the planning of training content with specific emphasis on individualized physical preparation relative to position and peak demands.

ACKNOWLEDGMENTS

We would like to thank Helene Davidsen Moa, Astrid G. Ree, Elias Koppen, Thomas Lyngbø, and Atle Ingebrigtsen for their assistance in data collection and the players in their respective clubs for approving to participate in the study. The study received funding from Tromsø Research Foundation and UiT The Arctic University of Norway.

CONFLICT OF INTEREST

No potential conflict of interest was reported by the authors.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

A. K. Winther https://orcid.org/0000-0001-5864-3445
M. B. Randers https://orcid.org/0000-0002-0192-8981
P. Krustrup https://orcid.org/0000-0002-1461-9838
S. A. Pettersen https://orcid.org/0000-0003-4700-0529

REFERENCES

1. FIFA. Physical analysis of the FIFA women's world cup France 2019. https://img.fifa.com/image/upload/zijqly4oednqa5gffgaz.pdf
2. Kirkendall DT. Evolution of soccer as a research topic. *Prog Cardiovasc Dis*. 2020;63(6):723-729. doi:10.1016/j.pcad.2020.06.011
3. Impellizzeri FM, Marcora SM, Coutts AJ. Internal and external training load: 15 years on. *Internat J Sports Physiol Performance*. 2019;14(2):270-273. doi:10.1123/ijspp.2018-0935
4. Mohr M, Krustrup P, Andersson H, Kirkendal D, Bangsbo J. Match activities of elite women soccer players at different performance levels. *J Strength Cond Res*. 2008;22(2):341-349. doi:10.1519/JSC.0b013e318165fef6
5. Gabbett TJ, Mulvey MJ. Time-motion analysis of small-sided training games and competition in elite women soccer players. *J Strength Cond Res*. 2008;22(2):543-552. doi:10.1519/JSC.0b013e3181635597
6. Scott D, Haigh J, Lovell R. Physical characteristics and match performances in women's international versus domestic-level football players: a 2-year, league-wide study. *Science Med Football*. 2020;4(3):211-215. doi:10.1080/24733938.2020.1745265
7. Andersson HA, Randers MB, Heiner-Møller A, Krustrup P, Mohr M. Elite female soccer players perform more high-intensity running when playing in international games compared with domestic league games. *J Strength Cond Res*. 2010;24(4):912-919. doi:10.1519/JSC.0b013e3181d09f21
8. Baptista I, Johansen D, Seabra A, Pettersen SA. Position specific player load during match-play in a professional football club. *PLoS One*. 2018;13(5). doi:10.1371/journal.pone.0198115
9. Schuth G, Carr G, Barnes C, Carling C, Bradley PS. Positional interchanges influence the physical and technical match performance variables of elite soccer players. *J Sports Sci*. 2016;34(6):501-508. doi:10.1080/02640414.2015.1127402
10. Bloomfield J, Polman R, O'Donoghue P. Physical demands of different positions in FA premier league soccer. *J Sports Sci Med*. 2007;6(1):63-70.
11. Bradley P, Di Mascio M, Peart D, Olsen P, Sheldon B. High-intensity activity profiles of elite soccer players at different performance levels. *J Strength Cond Res*. 2010;24(9):2343-2351. doi:10.1519/JSC.0b013e3181aeb1b3
12. Gabbett TJ, Wiig H, Spencer M. Repeated high-intensity running and sprinting in elite women's soccer competition. *Internat J Sports Physiol Performance*. 2013;8(2):130-138. doi:10.1123/ijspp.8.2.130
13. Datson N, Drust B, Weston M, Jarman IH, Lisboa PJ, Gregson W. Match physical performance of elite female soccer players during international competition. *J Strength Cond Res*. 2017;31(9):2379-2387. doi:10.1519/jsc.0000000000001575
14. Akenhead R, French D, Thompson KG, Hayes PR. The physiological consequences of acceleration during shuttle running. *Int J Sports Med*. 2015;36(4):302-307. doi:10.1055/s-0034-1389968
15. Gaudino P, Iaia FM, Alberti G, Strudwick AJ, Atkinson G, Gregson W. Monitoring training in elite soccer players: systematic bias between running speed and metabolic power data. *Int J Sports Med*. 2013;34(11):963-968. doi:10.1055/s-0033-1337943
16. Osgnach C, Poser S, Bernardini R, Rinaldo R, di Prampero PE. Energy cost and metabolic power in elite soccer: a new match

analysis approach. *Med Sci Sports Exerc*. 2010;42(1):170-178. doi:10.1249/MSS.0b013e3181ae5cfd
17. Mara JK, Thompson KG, Pumpa KL, Morgan S. The acceleration and deceleration profiles of elite female soccer players during competitive matches. *J Sci Med Sport*. 2017;20(9):867-872. doi:10.1016/j.jsams.2016.12.078
18. Panduro J, Ermidis G, Røddik L, et al. Physical performance and loading for six playing positions in elite female football: full-game, end-game, and peak periods. *Scand J Med Sci Sports*. 2021. doi:10.1111/sms.13877
19. Trewin J, Meylan C, Varley MC, Cronin J. The match-to-match variation of match-running in elite female soccer. *J Sci Med Sport*. 2018;21(2):196-201. doi:10.1016/j.jsams.2017.05.009
20. Strauss A, Sparks M, Pienaar C. The use of GPS analysis to quantify the internal and external match demands of semi-elite level female soccer players during a tournament. *J Sports Sci Med*. 2019;18(1):73-81.
21. Hewitt A, Norton K, Lyons K. Movement profiles of elite women soccer players during international matches and the effect of opposition's team ranking. *J Sports Sci*. 2014;32(20):1874-1880. doi:10.1080/02640414.2014.898854
22. Vescovi JD. Sprint profile of professional female soccer players during competitive matches: Female Athletes in Motion (FAiM) study. *J Sports Sci*. 2012;30(12):1259-1265. doi:10.1080/02640414.2012.701760
23. Vescovi JD, Favero TG. Motion characteristics of women's college soccer matches: female Athletes in Motion (FAiM) study. *Internat J Sports Physiol Performance*. 2014;9(3):405-414. doi:10.1123/ijspp.2013-0526
24. Barrett S, Midgley A, Reeves M, et al. The within-match patterns of locomotor efficiency during professional soccer match play: Implications for injury risk? *J Sci Med Sport*. 2016;19(10):810-815. doi:10.1016/j.jsams.2015.12.514
25. Gabbett TJ, Kennelly S, Sheehan J, et al. If overuse injury is a 'training load error', should undertraining be viewed the same way? *Br J Sports Med*. 2016;50(17):1017-1018. doi:10.1136/bjsports-2016-096308
26. Baptista I, Johansen D, Figueiredo P, Rebelo A, Pettersen SA. Positional differences in peak- and accumulated- training load relative to match load in elite football. *Sports*. 2019;8(1). doi:10.3390/sports8010001
27. Fereday K, Hills SP, Russell M, et al. A comparison of rolling averages versus discrete time epochs for assessing the worst-case scenario locomotor demands of professional soccer match-play. *J Sci Med Sport*. 2020;23(8):764-769. doi:10.1016/j.jsams.2020.01.002
28. Pettersen SA, Brenn T. Activity profiles by position in youth elite soccer players in official matches. *Sports Med Int Open*. 2019;3(1):E19-E24. doi:10.1055/a-0883-5540
29. Schimpchen J, Gopaladesikan S, Meyer T. The intermittent nature of player physical output in professional football matches: an analysis of sequences of peak intensity and associated fatigue responses. *Eur J Sport Sci*. 2020:1-10. doi:10.1080/17461391.2020.1776400
30. Martín-García A, Castellano J, Méndez Villanueva A, Gómez-Díaz A, Cos F, Casamichana D. Physical demands of ball possession games in relation to the most demanding passages of a competitive match. *J Sports Sci Med*. 2020;19(1):1-9.
31. Casamichana D, Castellano J, Diaz AG, Gabbett TJ, Martin-Garcia A. The most demanding passages of play in football competition: a comparison between halves. *Biol Sport*. 2019;36(3):233-240. doi:10.5114/biolsport.2019.86005
32. Oliva-Lozano JM, Rojas-Valverde D, Gómez-Carmona CD, Fortes V, Pino-Ortega J. Worst case scenario match analysis and contextual variables in professional soccer players: a longitudinal study. *Biol Sport*. 2020;37(4):429-436. doi:10.5114/biolsport.2020.97067
33. Bradley P, Noakes T. Match running performance fluctuations in elite soccer: indicative of fatigue, pacing or situational influences? *J Sports Sci*. 2013;31(15):1627-1638. doi:10.1080/02640414.2013.796062
34. Mohr M, Krustrup P, Bangsbo J. Fatigue in soccer: a brief review. *J Sports Sci*. 2005;23(6):593-599.
35. Fransson D, Krustrup P, Mohr M. Running intensity fluctuations indicate temporary performance decrement in top-class football. *Sci Med Football*. 2017;1(1):10-17.
36. Carling C, Dupont G. Are declines in physical performance associated with a reduction in skill-related performance during professional soccer match-play? *J Sports Sci*. 2011;29(1):63-71. doi:10.1080/02640414.2010.521945
37. 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. doi:10.1007/s40279-018-0965-6
38. Varley MC, Elias GP, Aughey RJ. Current match-analysis techniques' underestimation of intense periods of high-velocity running. *Internat J Sports Physiol Performance*. 2012;7(2):183-185. doi:10.1123/ijspp.7.2.183
39. Doncaster G, Page R, White P, Svenson R, Twist C. Analysis of physical demands during youth soccer match-play: considerations of sampling method and epoch length. *Res Q Exerc Sport*. 2020;91(2):326-334. doi:10.1080/02701367.2019.1669766
40. Beato M, Coratella G, Stiff A, Iacono AD. The validity and between-unit variability of GNSS units (STATSports Apex 10 and 18 Hz) for measuring distance and peak speed in team sports. *Front Physiol*. 2018;9:1288. doi:10.3389/fphys.2018.01288
41. Lozano D, Lampre M, Díez A, et al. Global positioning system analysis of physical demands in small and large-sided games with floaters and official matches in the process of return to play in high level soccer players. *Sensors*. 2020;20(22). doi:10.3390/s20226605
42. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol*. 2013;4:863.
43. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:14065823*. 2014.
44. Lenth R, Singmann H, Love J, Buerkner P, Herve M. Emmeans: estimated marginal means, aka least-squares means. R package version. 2018;1(1):3.
45. Ade JD, Drust B, Morgan OJ, Bradley PS. Physiological characteristics and acute fatigue associated with position-specific speed endurance soccer drills: production vs maintenance training. *Sci Med Football*. 2021;5(1):6-17. doi:10.1080/24733938.2020.1789202
46. Mohr M, Krustrup P. Comparison between two types of anaerobic speed endurance training in competitive soccer players. *J Human Kinetics*. 2016;51:183-192. doi:10.1515/hukin-2015-0181

47. Mohr M, Krustrup P, Bangsbo J. Fatigue in soccer: a brief review. *J Sports Sci*. 2005;23(6):593-599. doi:10.1080/02640410400021286
48. Ramos G, Nakamura F, Pereira L, et al. Movement patterns of a U-20 national women's soccer team during competitive matches: influence of playing position and performance in the first half. *Int J Sports Med*. 2017;38(10):747-754. doi:10.1055/s-0043-110767

How to cite this article: Winther AK, Baptista I, Pedersen S, et al. Position specific physical performance and running intensity fluctuations in elite women's football. *Scand J Med Sci Sports*. 2022;1(Suppl. 1):105–114. doi:10.1111/sms.14105

Paper II

Science and Medicine in Football

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rsmf20

The variability of physical match demands in elite women's football

Ivan Baptista, Andreas K. Winther, Dag Johansen, Morten B. Randers, Sigurd Pedersen & Svein Arne Pettersen

To cite this article: Ivan Baptista, Andreas K. Winther, Dag Johansen, Morten B. Randers, Sigurd Pedersen & Svein Arne Pettersen (2022) The variability of physical match demands in elite women's football, Science and Medicine in Football, 6:5, 559-565, DOI: 10.1080/24733938.2022.2027999

To link to this article: https://doi.org/10.1080/24733938.2022.2027999

SCIENCE AND MEDICINE IN FOOTBALL
2022, VOL. 6, NO. 5, 559–565
https://doi.org/10.1080/24733938.2022.2027999

Routledge
Taylor & Francis Group

ORIGINAL INVESTIGATIONS

OPEN ACCESS Check for updates

The variability of physical match demands in elite women's football

Ivan Baptista [a,b], Andreas K. Winther [c], Dag Johansen [a], Morten B. Randers [c,d], Sigurd Pedersen [c] and Svein Arne Pettersen [c]

[a]Department of Computer Science, Faculty of Science and Technology, UiT the Arctic University of Norway, Tromsø, Norway; [b]Faculty of Sport, University of Porto, Porto, Portugal; [c]School of Sport Sciences, Faculty of Health Sciences, UiT the Arctic University of Norway, Tromsø, Norway; [d]Department of Sports Science and Clinical Biomechanics, Sdu Sport and Health Sciences Cluster (Shsc), University of Southern Denmark, Odense, Denmark

ABSTRACT
Peak locomotor demands are considered as key metrics for conditioning drills prescription and training monitoring. However, research in female football has focused on absolute values when reporting match demands, leading to sparse information being provided regarding the degrees of variability of such metrics. Thus, the aims of this study were to investigate the sources of variability of match physical performance parameters in female football players and to provide a framework for the interpretation of meaningful changes between matches.
54 female players from four top-level clubs were monitored during one season. GPS APEX (STATSports, Northern Ireland), with a sampling frequency of 10 Hz, were used in 60 official matches (n = 393) to determine the full-match and 1-min peak locomotor demands of total distance (TD), high-speed running distance (HSRD), sprint distance (SpD), accelerations and decelerations (Acc/Dec) and peak speed (Pspeed). For each variable, the between-team, between-match, between-position, between-player, and within-player variability was estimated using linear mixed-effect modelling.
With exception to SpD (29.4 vs. 31.9%), all other metrics presented a higher observed match-to-match variability in the 1-min peaks than in the full-match (6.5 vs. 4.6%; 18.7% vs. 15.9%; 12.9 vs. 11.7%; for TD, HSRD and Acc/Dec, respectively). With the exception of SpD, higher changes in 1-min peaks than in full-match values are required to identify meaningful changes in each variable.
Different sources of variability seem to impact differently the match physical performance of female football players. Furthermore, to identify meaningful changes, higher changes in 1-min peaks than in full-match values are required.

ARTICLE HISTORY
Accepted 6 January 2022

KEYWORDS
Women's football; match-to-match variability; performance analysis; high-speed running; accelerations; peak locomotor demands

Introduction

The use of technology for monitoring match physical demands has become a common practice in professional football (Carling 2013). In recent years, the assessment of external load during official matches has evolved, partly due to the increasing prevalence of Global Positioning Systems (GPS) among football clubs (Whitehead et al. 2018), and the rule change in 2015 introduced by the International Football Association Board (IFAB) allowing the use of these technologies during official matches (FIFA 2015). Despite the growing body of knowledge within the match demands domain, the majority of the studies underestimate the true physical demands of competition, since several sport-specific movements (e.g., heading, tackling, accelerations and decelerations) are often omitted, leading to an under-estimation of match-load by 6–8% (Osgnach et al. 2010). The detailed performance data obtained through the analysis of match running activity and acceleration metrics can be used by practitioners to profile the player's game requirements and consequently guide decision-making throughout the microcycle, such as the adjustment of recovery sessions or to establish physical targets during the week (Al Haddad et al. 2018).

Football performance is a multifactorial construct with a dynamic and stochastic nature, where players' physical performances (e.g., high-speed activities) are affected by external factors (e.g., ball possession and period of the season) which consequently causes a fluctuation of these metrics between consecutive matches (Gregson et al. 2010). The variability in a football player's performance from match to match can provide estimates of the smallest worthwhile change, an important piece of information for sport scientists monitoring players or for scientists designing and analysing studies on factors affecting performance (Hopkins et al. 1999). This concept has been deeply studied in men's football (Bush et al. 2015; Carling et al. 2016; Gonçalves et al. 2018; Oliva-Lozano et al. 2021) and demonstrated by the coefficient of variation (CV) of a particular physical performance parameter (Novak et al. 2021). Previous studies have shown that this match-to-match variability can be caused by internal (e.g., fitness characteristics) and external factors (ball possession in match-play) (Carling et al. 2016) including the method used for match analysis (Randers et al. 2010; Pettersen et al. 2018). Previous research in men's football has been unanimous when reporting high-speed running as the most inconsistent variable from match-to-match (Bush et al. 2015; Carling et al. 2016; Trewin et al. 2017)

CONTACT Ivan Baptista ivantm@fade.up.pt 205 1D, 4200-104 Porto, Portugal

with Gregson et al. (Gregson et al. 2010) adding that this variability (CV~15% to 30%) is higher for central positions than for wide positions.

However, little research has been done within this field in the women's football context. Although both men and women play the same game, research in other sports, such as weightlifting (McGuigan and Kane 2004) and cycling (Paton and Hopkins 2006) has shown a tendency for greater variability in women compared to men. In a recent study of a women's national team, Trewin et al. (Trewin et al. 2018) reported a higher occurrence and lower variability of accelerations (CV = 17%), when compared to high-speed running and sprint efforts (CV = 34% and 56%, respectively). These results are in line with research in men's football (Dalen et al. 2019) where accelerations have been proposed to be a more stable and sensitive measure of physical performance than high-speed running activities. The study of female national team players (Trewin et al. 2018) also presented the high-speed running and sprint efforts of centre backs (CB) as the metrics with the greatest variation when compared to other playing positions (CV = 41–65%). Despite the novelty of the study, Trewin et al. (Trewin et al. 2018) analysed data from a single national team, across five consecutive seasons, which should be considered as a possible bias of the results, since within this time span changes in the physical condition of the players are very likely to occur (Mohr et al. 2003). Moreover, a multiple team analysis would be beneficial in order to reduce the possible bias caused by certain contextual factors (i.e., team) in the match-to-match variability observed.

Research within the match analysis domain is no longer bound to the analysis of absolute (full match) values, and the concept of peak locomotor demands has been gathering researchers' attention over the last years (Weaving et al. 2019). Previous research has suggested that match average demands are not the most informative outcomes for players preparation, since the use of such values to characterize match physical demands will most likely underestimate the most intense periods of the match (Delaney et al. 2015). Although, more common terms, such as peak period (Baptista et al. 2019a, most demanding passages of play (Martin-Garcia et al. 2018; Castellano et al. 2020) and worst-case scenarios (Cunningham et al. 2018; Fereday et al. 2020) have been used to refer to this concept. Researchers and practitioners should also be aware that only univariate locomotor measurements have been presented and that such an approach does not represent the total amount of activity (Novak et al. 2021). Therefore, to minimize such misinterpretation of the concept, this paper will use the term suggested by Weaving et al. (Weaving et al. 2019) and further supported by Novak et al. (Novak et al. 2021) – *peak locomotor demands*. Despite the growing interest in studying the training and match demands in female football (Gabbett and Mulvey 2008; Mohr et al. 2008; Andersson et al. 2010; Vescovi 2012; Gabbett et al. 2013; Hewitt et al. 2014; Vescovi and Favero 2014; Datson et al. 2017; Mara et al. 2017; Vescovi and Falenchuk 2019), this representation of external load has focused on absolute values (full-match) or long fixed-periods (i.e., 15 minutes), with sparse information provided about shorter peak locomotor demands (e.g., 1, 3 or 5 minutes) of female competitions(Trewin et al. 2018; Harkness-Armstrong et al. 2020; Panduro et al. 2021). This can in turn lead to limited information for training prescription, since peak locomotor demands have been suggested as key-metrics for the prescription of conditioning drills and the monitoring of training intensities.(Whitehead et al. 2018)

Irrespectively, the random factors (i.e., match, position, players, and team) become important to determine the different degrees of variability of key physical variables, so practitioners can make more evidence-based decisions in their daily practices. Quantifying the match-to-match variability of different physical variables may be used to determine whether a change in match demands can be considered as normal or unusual (Oliva-Lozano et al. 2021). Therefore, the aim of this study was twofold: 1) to investigate the different sources of variability of selected match physical performance parameters in elite football player cohorts, using full match values and 1-min peak locomotor demands; and 2) to provide reference values for interpreting changes in match physical performance.

Methods

Participants and match samples

With ethical institutional approval and written informed consent from the participants, 108 female football players (22.4 ± 4.0 years of age) from four elite-level (top tier division) Norwegian clubs participated in the study. Player movement data from one season (2020) including 60 official matches was collected using GPS APEX (STATSports, Northern Ireland), with a sampling frequency of 10 Hz. The validity and acceptable levels of accuracy (bias <5%) of this tracking system have previously been presented (Beato et al. 2018). During matches, each player wore a tight vest with a GPS unit on the back of their upper body between scapula as described by the manufacturer. The microsensor devices were activated 15 min before the start of each match, in accordance with the manufacturer's recommendations and previous research (Lozano et al. 2020), with this period of time excluded from analyses. To minimize inter-devices error (Beato et al. 2018), each player used the same GPS unit for the entire season. The mean number of satellites and horizontal dilution of precision was 17.5 ± 2.8 and 1.4 ± 0.6, respectively.

Data processing

Doppler derived speed data were exported from manufacturer software (STATSport Sonra 2.1.4) into Python 3.7.6. for processing (linearly interpolating any missing raw data), and to derive metrics. Raw acceleration was then calculated over a period of 0.6 seconds. Matches were treated in which two of our teams played against each other as separate matches, and, because of positional differences in locomotor demands, the same player in a new position was treated as a new player. Goalkeepers were excluded from analysis and the selected playing positions, (central defenders, full-backs, midfielders, wide midfielders, and forwards), were chosen according to previous research.(Schuth et al. 2016; Baptista et al. 2018) To get a representative sample, players were included only if: a) completed, at least, two full-time (90 min) matches; b) and played the entire match in the same playing position. Match performance data of <90 min was treated as missing. This resulted in an initial sample of 501 observations with 108 missing values, which were subsequently removed in the complete case analysis. The final sample included 393 match observations (M_{obs}) from 54 players (central defenders, n = 10, M_{obs} = 113; full-backs, n = 11, M_{obs} = 84; central midfielders, n = 16, M_{obs} = 105; wide midfielders, n = 9; M_{obs} = 57 and central forwards, n = 8, M_{obs} = 34).

Physical performance variables

The physical parameters analysed included: total distance (TD), high-speed running distance (HSRD) (>4.44 $m\cdot s^{-1}$), sprint distance (SpD) (>5.55 $m\cdot s^{-1}$), number of accelerations and decelerations (Acc/Dec), and peak speed (Pspeed). In accordance with Trewin et al. (Trewin et al. 2017), accelerations and decelerations were defined as a positive or negative change in speed of more than ±2.26 $m\cdot s^{-2}$, with a minimal effort duration of 0.3 seconds, finishing when the rate of acceleration/deceleration reached 0 $m\cdot s^{-2}$. The speed thresholds were chosen according to previous research.(Trewin et al. 2018; Strauss et al. 2019) Except for Pspeed, all other variables were used to analyse both full match (absolute values) and peak locomotor demands (1-min rolling analysis period). The epoch length for the peak locomotor demands was chosen according to the findings of Doncaster et al. (Doncaster et al. 2020), where 1 min epochs produced the highest relative intensities when compared with 3- and 5-min epochs.

Statistical Methods

After deriving all the metrics, the data were transferred to R (R.4.0.5, R Core Team, 2021) for statistical analysis. To estimate the sources of variability (between-team, between-position, between-player, between-match, and the residual within-player variability) and to provide reference values for interpreting changes in match physical performance, we used a similar approach as Oliva-Lozano et al. (Gonçalves et al. 2018). The design located units of analysis (individual match observations) nested within clusters of units (players), further nested within playing positions and teams. To account for this hierarchical (correlated) nesting, and to quantify the variability in match physical performance, data were analysed using linear mixed-effect modelling with the package lme4 (Bates et al. 2015). For each physical parameter, the model was specified to include a random intercept for the random effects: team, position, player ID, and match ID. All models were estimated via Restricted Estimated Maximum Likelihood (REML), and model appropriateness was verified by examining the QQ-plots of the studentized residuals. Each random effect represented a source of variability and was expressed in raw units (standard deviation – SD) by modelling the original data, and in percentage units (CV%) by first log-transforming the original data before modelling, and then back-transforming each estimate after modelling was done (Hopkins et al. 2009).

Similar to Oliva-Lozano et al. (Oliva-Lozano et al. 2021), variability estimates were used to provide a framework for practitioners to interpret individual changes in indicators of match physical performance. Here, 80% and 90% limits of agreement (LoA) were calculated by multiplying the square root of 2 with the appropriate values from the t-distribution (with infinite degrees of freedom) and the observed between-match variability expressed (e.g., the pooled between-match and within-player variability). Furthermore, practical significant changes associated with alpha levels of 0.10 and 0.05 were calculated using the formula: * observed between-match variability * t-statistic + threshold. Here, the *observed between-match variability* was the same as described above, while the *threshold* term was equivalent to the smallest worthwhile change (0.2 * the observed between-player variability – or the pooled between-player and within-player variability).

Table 1. Variability of full match and 1-min peak locomotor demands expressed in raw units and coefficients of variation (%).

	Metric		Variability				
			Between-match	Between-team	Between-position	Between-player	Within-player
SD (90% CI)[a]	TD (m)	Full match	335 (278–393)	37 (0–212)	456 (132–749)	473 (379–547)	259 (239–277)
		1′ peak	6 (4–7)	2 (0–4)	8 (2–14)	7 (5–8)	10 (10–11)
	HSRD (m)	Full match	132 (103–154)	51 (0–137)	288 (95–446)	272 (222–323)	160 (148–171)
		1′ peak	7 (5–8)	2 (0–6)	10 (2–15)	8 (6–10)	13 (12–14)
	SpD (m)	Full match	40 (29–49)	0 (0–39)	111 (31–172)	103 (84–122)	73 (68–78)
		1′ peak	3 (0–4)	1 (0–3)	8 (3–12)	6 (4–7)	11 (10–12)
	Acc/Dec (#)	Full match	12 (10–15)	0 (0–11)	19 (0–32)	28 (23–33)	20 (18–21)
		1′ peak	0.3 (0.0–0.4)	0.0 (0.0–0.3)	0.8 (0.2–1.2)	0.7 (0.5–0.9)	1.3 (1.2–1.3)
	$Peak_{speed}$ (m/s)	Full match	0.1 (0.0–0.1)	0.0 (0.0–0.1)	0.1 (0.0–0.9)	0.3 (0.3–0.4)	0.3 (0.3–0.3)
CV (90% CI)[b]	TD (m)	Full match	3.6 (3.0–4.2)	0.2 (0.0–2.2)	4.9 (1.7–8.0)	4.9 (4.0–5.8)	2.8 (2.6–3.0)
		1′ peak	3.1 (2.3–3.8)	0.7 (0.0–2.2)	4.7 (1.5–7.3)	3.7 (2.7–4.5)	5.7 (5.3–6.1)
	HSRD (m)	Full match	10.2 (8.1–12.2)	1.1 (0.0–7.3)	22.8 (7.1–37.4)	18.9 (15.0–22.8)	11.7 (10.9–12.5)
		1′ peak	7.6 (5.5–9.7)	1.4 (0.0–5.2)	12.8 (2.6–20.9)	10.7 (7.9–13.3)	16.7 (15.5–17.9)
	SpD (m)	Full match	13.8 (10.0–17.5)	0.0 (0.0–13.1)	39.3 (8.8–66.9)	37.2 (28.4–46.2)	27.7 (25.6–29.7)
		1′ peak	6.4 (0.0–9.7)	0.0 (0.0–7.2)	20.0 (6.6–32.1)	14.9 (10.0–19.0)	28.4 (26.4–30.6)
	Acc/Dec (#)	Full match	6.2 (4.7–7.6)	0.0 (0.0–6.0)	9.2 (0.0–15.6)	14.2 (11.0–17.0)	9.7 (9.1–10.4)
		1′ peak	0.3 (0.0–0.4)	0.0 (0.0–3.2)	7.1 (0.8–11.3)	7.3 (5.1–9.0)	12.6 (11.7–13.5)
	$Peak_{speed}$ (m/s)	Full match	0.1 (0.0–0.1)	0.0 (0.0–1.7)	1.9 (0.0–3.7)	4.4 (3.4–5.2)	4.4 (4.0–4.7)

SD = Standard deviation; CI = Confidence Intervals; CV = Coefficient of variation.
[a]Values presented in the metric's unit of measurement;
[b]Values presented as a percentage of the mean

Results

The decomposed variability of full match and 1-min peak match analysis metrics are presented in Table 1. All estimates of between-position, between-match, between-player, within-player, and between-team are expressed in raw (SD) and percentage (CV) units. CV values of full match variables ranged from 0.0% to 39.3%, with the lowest CVs associated with between-team variability of Pspeed (0.0%) and the highest with between-position variability of SpD (39.3%). With the exception of between-team variability, which presented low values for all metrics, all sources of variability of full match metrics were greater for SpD (13–39%) when compared with all other external load variables. Between-player (for TD, Acc/Dec and Pspeed) and between-position analysis (for HSRD and SpD) present higher CVs, in the full match variables analysed, relative to the other sources of variability. CV values of 1-min peak variables ranged from 0.0% to 28.4%, with the lowest CVs associated with the between-team variability of Acc/Dec (0.0%) and the highest with the within-player variability of SpD (28.4%). The within-player variability assumes the largest CVs for the 1-min peak variables.

The observed match-to-match variability (combined between-match and within-player) and reference values for interpreting individual changes are presented in Table 2. With exception to SpD (29.4 vs. 31.9%), all other metrics presented a higher observed match-to-match variability in the 1-min peaks than in the full match (6.5 vs. 4.6%; 18.7% vs. 15.9%; 12.9 vs. 11.7%; for TD, HSRD and Acc/Dec, respectively). Based on the model used to identify significant changes (see methods section), between-match individual changes of ±9% (α = 0.10) and ±12% (α = 0.05) in full match metrics of TD and Pspeed would be considered unusual and suggest practical significance. For HSRD (33%; 42%), SpD (68%; 84%) and Acc/Dec (25%; 31%) these thresholds (α = 0.10; α = 0.05; respectively) are considerably higher. Regarding 1-min peaks, and with exception to SpD, higher changes than in full-match values are required to identify meaningful difference.

Discussion

Full-match vs. 1-min peak variability

This study is novel, being the first that decomposes and compares the variability of absolute (full-match) and relative (1-min peak) match external load metrics in elite women's football. A novel finding was the higher observed match-to-match variability in 1-min peaks when compared to the full match, in TD (6.5% vs. 4.6%), HSRD (18.7% vs. 15.9%) and Acc/Dec (12.9% vs. 11.7%). This difference may be caused by external factors (e.g., match result and opponent) alongside the dynamic and stochastic nature of a football match, which in this case seasonal fluctuations appear to have had a higher influence in the most demanding periods than in the mean match values.(Gregson et al. 2010) While not having reference to female football, previous research in male football (Novak et al. 2021) presented CV values of 3-min peaks similar to our study, for TD (~7%), HSRD (~21-31%) and SpD (~35-56%). This information is particularly relevant since the study of univariate peak locomotor demands has been used by practitioners to inform training prescription (Baptista et al. 2019a), and consequently as a strategy to better prepare their players to cope with these peaks during match-play. However, as previously observed in absolute values (Carling et al. 2016), peak locomotor demands are also unstable across matches. The poor consistency of specific peak high-speed metrics presented in men's football (Novak et al. 2021), and here corroborated for women's football, may raise questions regarding its practical applicability. Although the analysis of peak locomotor demands in matches has become a common trend among practitioners, its applicability as benchmarks for training sessions may be controversial.

Sources of variability

After decomposing the variability into five different sources (between-match, between-position, between-player, within-player and between-team), we observed that all sources were greater for SpD than for the other physical metrics, both in full match (13.8–39.3%) and 1-min peaks (6.4–28.4%), with a minor exception in the between-team variability, where HSRD (~1%) presented slightly higher CV than SpD (~0%). These results are in line with previous research in male football, where the highest CV values were observed in high-speed metrics (Gregson et al. 2010;

Table 2. Reference values for interpretation of individual changes in match physical performance in full match and 1-min peak periods.

Metric		Observed match-to-match variability[b] CV (90% CI)[a]	± Limits of agreement (%)[c] 80%	90%	Change (±) required to be practically significant (%)[c] α = 0.10	α = 0.05
TD	Full match	4.6 (4.1–5.0)	8.3	10.6	9.4	11.8
	1′ peak	6.5 (6.0–7.0)	11.8	15.1	13.2	16.5
HSRD	Full match	15.9 (14.3–17.7)	28.8	36.9	33.3	41.5
	1′ peak	18.7 (17.2–20.1)	33.9	43.5	38.0	47.6
SpD	Full match	31.9 (29.1–34.5)	57.7	74.1	67.5	83.9
	1′ peak	29.4 (27.1–31.6)	53.3	68.4	59.9	75.0
Acc/Dec	Full match	11.7 (10.8–12.6)	21.2	27.2	24.7	30.7
	1′ peak	12.9 (12.1–13.8)	23.4	30.0	26.3	33.0
Pspeed	Full match	4.5 (4.2–4.8)	8.1	10.4	9.4	11.7

CV = Coefficient of variation.
[a]Values presented as a percentage of the mean.
[b]Based on the combined between-match and within-player variability.
[c]Based on the observed match-to-match variability

Carling et al. 2016). For instance, Carling et al. (Carling et al. 2016) presented greater variability for distances above 7.0 $m{\cdot}s^{-1}$ (37%) than for distances between 5.5 and 7.0 $m{\cdot}s^{-1}$ (18.1%). These discrepancies between locomotor categories (full-match values) are somewhat similar to those presented in our study, where the observed match-to-match variability of SpD (31.9%) presented twice the magnitude of HSRD (15.9%).

Using a similar approach of previous research (Oliva-Lozano et al. 2021), we separately analysed the elements occurring at the match and player level by partitioning the observed match-to-match variability into between-match and within-player variability. Our full match results for TD (3.6% vs. 2.8%) and Pspeed (1.0% vs. 4.4%) were identical to those reported by Oliva-Lozano et al. (Oliva-Lozano et al. 2021) (4.3% vs. 3.7% and 1.5% vs. 4.9%; for TD and Pspeed, respectively), where these metrics appeared relatively stable both for between-match and within-player variability. However, regarding Acc/Dec our study presents a lower CV for between-match than for between-position variability (6.2% vs. 9.7%), while the study of Oliva-Lozano et al. (Oliva-Lozano et al. 2021) reported an opposite trend (4.9% vs. 2.6%). We conjecture that the presence of a high between-position variability could be caused by the divergent individual characteristics within the playing position. In fact, our study presented a higher sample size, and consequently more players per position than the study of Oliva-Lozano et al. (Oliva-Lozano et al. 2021), meaning the presence of a larger diversity of players within each position. Furthermore, the between-match (10.2%) and within-player (11.7%) variability of HSRD observed in our study were considerably lower than reported in men's teams (19% and 23%, respectively) (Oliva-Lozano et al. 2021). We conjecture that this discrepancy between studies is caused by the different high-speed running thresholds used in female (>4.44 $m{\cdot}s^{-1}$) and male teams (>5.8 $m{\cdot}s^{-1}$), which is associated with the fact that variability tends to increase with running intensity,(Carling et al. 2016) justifies such differences.

Individual changes interpretation

By partitioning the match physical performance variability into different sources, we provide valuable information that may assist football coaches to make more evidence-based decisions regarding the monitoring of between-match changes. The reference values for interpreting the individual changes presented in Table 2 were obtained by a combination of between-match and within-player variability, resulting in 80% and 90% LoA, which were then complemented with thresholds for practical significance (see Methods section). For example, according to our results, a player's positive or negative variation in the match Pspeed of >9.4% (α = 0.10) should be considered unusual, while a change in HSRD peak period of <47.6% (α = 0.05) could be interpreted as usual. Previous research(Stevens et al. 2017; Baptista et al. 2019a) have suggested that the interpretation of training load data is facilitated if match load is used as a reference, allowing a more appropriate training prescription and communication between practitioners. Therefore, understanding the meaningfulness and practical significance of match physical performance variability may help coaches during the training load management process. For instance, a marked decrease in HSRD from one match to another does not necessarily mean a lower physical condition of the player. Consequently, before making hasty conclusions, practitioners may firstly confirm if such variation falls within the practical significant range.

Limitations and further research

Following the suggestion of Oliva-Lozano et al. (Oliva-Lozano et al. 2021) for the necessity to conduct a multi-club study, we included four different top-level teams. This strategy has the added benefit of likely increasing the data heterogeneity and consequently diminishing the risk of bias caused by a specific style of play and/or training periodization (Baptista et al. 2019b). However, the low values for between-team variability may suggest that our data contain too few and too homogenous clusters. Future studies should try to remedy this by including more teams from a broader range of performance level within a division. Other limitations include the fact that GPS may present lower accuracy than radio-based local positioning systems (Pettersen et al. 2018), particularly for high-speed measures like HSRD, SpD and Pspeed (Buchheit and Simpson 2017). We also recognize that positional differences will likely affect the magnitude of the variability and thus, future research should also attempt to present results by playing position. Despite the deliberate exclusion of the warm-up data, at a finer granularity, this pre-match period might influence the players' readiness and preparedness for the game. Furthermore, in this study only univariate peak locomotor demands were considered and, therefore, different conclusions could be drawn if multivariate peak periods were analysed.

Conclusion

In general, match physical performance of female football players seems to be affected differently by the different sources of variability. Moreover, the high-speed metrics presented a higher observed match-to-match variability than the other key-metrics analysed. Finally, higher changes in 1-min peaks than in full-match values are required to be considered meaningful. The outcomes of the present study may address reference values that allow coaches to better interpret the inevitable variation of match physical performance. Practitioners must consider performance variability as advantageous and keep in mind that such a phenomenon is part of the team sports nature. Therefore, training prescription should avoid using specific benchmarks to achieve, but rather promote the presence of varied training stimulus and intensities, as well as use reference values for interpreting individual changes in match physical performances.

Acknowledgements

The authors would like to thank Astrid Grøttå Ree, Helene Davidsen Moa, Thomas Lyngbø, Mads Pettersen, Elias Koppen, and Atle Ingebrigtsen for their help in the data collection.

Author contributions

Conceptualization, IB, AKW, MBR and SAP; Data curation, AKW; Formal analysis, AKW; Investigation, IB and AKW; Methodology, IB, AKW, DJ, MBR and SAP; Project administration, SAP; Supervision, DJ and SAP; Writing—original draft, IB; Writing—review & editing, AKW, DJ, MBR, SP and SAP. All authors have read and agreed to the published version of the manuscript.

Data availability statement

The data that support the findings of this study are available upon reasonable request.

Disclosure statement

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

Funding

This research was funded by Tromsø Research Foundation and UiT, The Arctic University of Norway. The funding body had no role in designing the study and data collection, analysis, interpretation, and in writing of the study. There is no grant/award number to report;Tromsø Research Foundation and UiT, The Arctic University of Norway.

ORCID

Ivan Baptista http://orcid.org/0000-0002-0330-7636
Morten B. Randers http://orcid.org/0000-0002-0192-8981
Svein Arne Pettersen http://orcid.org/0000-0003-4700-0529

References

Al Haddad H, Mendez-Villanueva A, Torreno N, Munguía-Izquierdo D, Suárez-Arrones L. 2018. Variability of GPS-derived running performance during official matches in elite professional soccer players. J Sports Med Phys Fitness. 58(10):1439–1445. doi:10.23736/S0022-4707.17.07500-4.

Andersson HA, Randers MB, Heiner-Møller A, Krustrup P, Mohr M. 2010. Elite female soccer players perform more high-intensity running when playing in international games compared with domestic league games. J Strength Cond Res. 24(4):912–919. doi:10.1519/JSC.0b013e3181d09f21.

Baptista I, Johansen D, Figueiredo P, Rebelo A, Pettersen SA. 2019a. Positional differences in peak- and accumulated- training load relative to match load in elite football. Sports (Basel, Switzerland). 8(1). doi:10.3390/sports8010001.

Baptista I, Johansen D, Figueiredo P, Rebelo A, Pettersen SA. 2019b. A comparison of match-physical demands between different tactical systems: 1-4-5-1 vs 1-3-5-2. PLoS One. 14(4):e0214952. doi:10.1371/journal.pone.0214952.

Baptista I, Johansen D, Seabra A, Pettersen SA. 2018. Position specific player load during match-play in a professional football club. PLoS One. 13(5). doi:10.1371/journal.pone.0198115.

Bates D, Mächler M, Bolker B, Walker, S . 2015. Fitting linear mixed-effects models using lme4 Journal of Statistical Software 67 (1): :1–48 doi:10.18637/jss.v067.i01

Beato M, Coratella G, Stiff A, Iacono AD. 2018. The validity and between-unit variability of GNSS units (STATSports apex 10 and 18 Hz) for measuring distance and peak speed in team sports. Front Physiol. 9:1288. doi:10.3389/fphys.2018.01288.

Buchheit M, Simpson BM. 2017. Player-tracking technology: half-full or half-empty glass?. Int J Sports Physiol Perform. 12(Suppl 2):S235–s41. doi:10.1123/ijspp.2016-0499.

Bush MD, Archer DT, Hogg R, Bradley PS. 2015. Factors influencing physical and technical variability in the English Premier League. Int J Sports Physiol Perform. 10(7):865–872. doi:10.1123/ijspp.2014-0484.

Carling C. 2013. Interpreting physical performance in professional soccer match-play: should we be more pragmatic in our approach?. Sports Med (Auckland, NZ). 43(8):655–663. doi:10.1007/s40279-013-0055-8.

Carling C, Bradley P, McCall A, Dupont G. 2016. Match-to-match variability in high-speed running activity in a professional soccer team. J Sports Sci. 34(24):2215–2223. doi:10.1080/02640414.2016.1176228.

Castellano J, Martin-Garcia A, Casamichana D. 2020. Most running demand passages of match play in youth soccer congestion period. Biol Sport. 37 (4):367–373. doi:10.5114/biolsport.2020.96853.

Cunningham DJ, Shearer DA, Carter N, Drawer S, Pollard B, Bennett M, Eager R, Cook CJ, Farrell J, Russell M, et al. 2018. Assessing worst case scenarios in movement demands derived from global positioning systems during international rugby union matches: rolling averages versus fixed length epochs. PLoS One. 13(4):e0195197. doi:10.1371/journal.pone.0195197.

Dalen T, Lorås H, Hjelde GH, Kjøsnes TN, Wisløff U. 2019. Accelerations - a new approach to quantify physical performance decline in male elite soccer?. Eur J Sport Sci. 19(8):1015–1023. doi:10.1080/17461391.2019.1566403.

Datson N, Drust B, Weston M, Jarman IH, Lisboa PJ, Gregson W. 2017. Match physical performance of elite female soccer players during international competition. J Strength Cond Res. 31(9):2379–2387. doi:10.1519/JSC.0000000000001575.

Delaney JA, Scott TJ, Thornton HR, Bennett KJM, Gay D, Duthie GM, Dascombe BJ. 2015. Establishing duration-specific running intensities from match-play analysis in rugby league. Int J Sports Physiol Perform. 10(6):725–731. doi:10.1123/ijspp.2015-0092.

Doncaster G, Page R, White P, Svenson R, Twist C. 2020. Analysis of physical demands during youth soccer match-play: considerations of sampling method and epoch length. Res Q Exerc Sport. 91 (2):326–334. doi:10.1080/02701367.2019.1669766.

Fereday K, Hills SP, Russell M, Smith J, Cunningham DJ, Shearer D, McNarry M, Kilduff LP. 2020. A comparison of rolling averages versus discrete time epochs for assessing the worst-case scenario locomotor demands of professional soccer match-play. J Sci Med Sport. 23 (8):764–769. doi:10.1016/j.jsams.2020.01.002.

FIFA. 2015. FIFA and IFAB to develop global standard for electronic performance and tracking systems Retrieved 13 february 2021 [Available from: https://www.fifa.com/who-we-are/news/fifa-and-ifab-to-develop-global-standard-for-electronic-performance-an-2709918.

Gabbett TJ, Mulvey MJ. 2008. Time-motion analysis of small-sided training games and competition in elite women soccer players. J Strength Cond Res. 22(2):543–552. doi:10.1519/JSC.0b013e3181635597.

Gabbett TJ, Wiig H, Spencer M. 2013. Repeated high-intensity running and sprinting in elite women's soccer competition. Int J Sports Physiol Perform. 8(2):130–138. doi:10.1123/ijspp.8.2.130.

Gonçalves B, Coutinho D, Travassos B, Folgado H, Caixinha P, Sampaio J. 2018. Speed synchronization, physical workload and match-to-match performance variation of elite football players. PLoS One. 13(7): e0200019. doi:10.1371/journal.pone.0200019.

Gregson W, Drust B, Atkinson G, Salvo V. 2010. Match-to-match variability of high-speed activities in premier league soccer. Int J Sports Med. 31(4):237–242. doi:10.1055/s-0030-1247546.

Harkness-Armstrong A, Till K, Datson N, Emmonds, S . 2020. Whole and peak physical characteristics of elite youth female soccer match-play. J Sports Sci 39 (12) :1320–1329.

Hewitt A, Norton K, Lyons K. 2014. Movement profiles of elite women soccer players during international matches and the effect of opposition's team ranking. J Sports Sci. 32(20):1874–1880. doi:10.1080/02640414.2014.898854.

Hopkins WG, Hawley JA, Burke LM. 1999. Design and analysis of research on sport performance enhancement. Med Sci Sports Exerc. 31(3):472–485. doi:10.1097/00005768-199903000-00018.

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

Lozano D, Lampre M, Díez A, Gonzalo-Skok O, Jaén-Carrillo D, Castillo D, Arjol JL. 2020. Global positioning system analysis of physical demands in small and large-sided games with floaters and official matches in the process of return to play in high level soccer players. Sens (Basel). 20(22). doi:10.3390/s20226605.

Mara JK, Thompson KG, Pumpa KL, Morgan S. 2017. The acceleration and deceleration profiles of elite female soccer players during competitive matches. J Sci Med Sport. 20(9):867–872. doi:10.1016/j.jsams.2016.12.078.

Martin-Garcia A, Casamichana D, Diaz AG, Cos, F, Gabbett, TJ . 2018. Positional differences in the most demanding passages of play in football competition. J Sports Sci Med. 17(4):563–570.

McGuigan MR, Kane MK. 2004. Reliability of performance of elite Olympic weightlifters. J Strength Cond Res. 18(3):650–653. doi:10.1519/12312.1.

Mohr M, Krustrup P, Andersson H, Kirkendal D, Bangsbo J. 2008. Match activities of elite women soccer players at different performance levels. J Strength Cond Res. 22(2):341–349. doi:10.1519/JSC.0b013e318165fef6.

Mohr M, Krustrup P, Bangsbo J. 2003. Match performance of high-standard soccer players with special reference to development of fatigue. J Sports Sci. 21(7):519–528. doi:10.1080/0264041031000071182.

Novak AR, Impellizzeri FM, Trivedi A, Rennie MJ, Watsford ML, Fransen J. 2021. Analysis of the worst-case scenarios in an elite football team: towards a better understanding and application. J Sports Sci. 1–10. doi:10.1080/02640414.2021.1995245.

Oliva-Lozano JM, Muyor J, Fortes V, McLaren, SJ. 2021. Decomposing the variability of match physical performance in professional soccer: implications for monitoring individuals. Eur J Sport Sci 21 (11) :1588–1596.

Osgnach C, Poser S, Bernardini R, Rinaldo R, Di Prampero PE. 2010. Energy cost and metabolic power in elite soccer: a new match analysis approach. Med Sci Sports Exerc. 42(1):170–178. doi:10.1249/MSS.0b013e3181ae5cfd.

Panduro J, Ermidis G, Røddik L, Vigh-Larsen JF, Madsen EE, Larsen MN, Pettersen SA, Krustrup P, Randers MB. 2021. Physical performance and loading for six playing positions in elite female football: full-game, end-game, and peak periods. Scand J Med Sci Sports. doi:10.1111/sms.13877.

Paton CD, Hopkins WG. 2006. Variation in performance of elite cyclists from race to race. Eur J Sport Sci. 6(1):25–31. doi:10.1080/17461390500422796.

Pettersen SA, Johansen HD, Baptista I, Halvorsen P, Johansen D. 2018. Quantified soccer using positional data: a case study. Front Physiol. 9 (866). doi:10.3389/fphys.2018.00866.

Randers MB, Mujika I, Hewitt A, Santisteban J, Bischoff R, Solano R, Zubillaga A, Peltola E, Krustrup P, Mohr M, et al. 2010. Application of four different football match analysis systems: a comparative study. J Sports Sci. 28(2):171–182. doi:10.1080/02640410903428525.

Schuth G, Carr G, Barnes C, Carling C, Bradley PS. 2016. Positional interchanges influence the physical and technical match performance variables of elite soccer players. J Sports Sci. 34(6):501–508. doi:10.1080/02640414.2015.1127402.

Stevens TGA, de Ruiter CJ, Twisk JWR, Savelsbergh GJP, Beek PJ. 2017. Quantification of in-season training load relative to match load in professional Dutch Eredivisie football players. Sci Med Footb. 1 (2):117–125. doi:10.1080/24733938.2017.1282163.

Strauss A, Sparks M, Pienaar C. 2019. The use of GPS analysis to quantify the internal and external match demands of semi-elite level female soccer players during a tournament. J Sports Sci Med. 18(1):73–81.

Trewin J, Meylan C, Varley MC, Cronin J. 2017. The influence of situational and environmental factors on match-running in soccer: a systematic review. Sci Med Footb. 1(2):183–194. doi:10.1080/24733938.2017.1329589.

Trewin J, Meylan C, Varley MC, Cronin J. 2018. The match-to-match variation of match-running in elite female soccer. J Sci Med Sport. 21(2):196–201. doi:10.1016/j.jsams.2017.05.009.

Vescovi JD. 2012. Sprint profile of professional female soccer players during competitive matches: female athletes in motion (FAiM) study. J Sports Sci. 30(12):1259–1265. doi:10.1080/02640414.2012.701760.

Vescovi JD, Falenchuk O. 2019. Contextual factors on physical demands in professional women's soccer: female athletes in motion study. Eur J Sport Sci. 19(2):141–146. doi:10.1080/17461391.2018.1491628.

Vescovi JD, Favero TG. 2014. Motion characteristics of women's college soccer matches: female athletes in motion (FAiM) study. Int J Sports Physiol Perform. 9(3):405–414. doi:10.1123/ijspp.2013-0526.

Weaving D, Sawczuk T, Williams S, Scott T, Till K, Beggs C, Johnston RD, Jones B. 2019. The peak duration-specific locomotor demands and concurrent collision frequencies of European Super League rugby. J Sports Sci. 37(3):322–330. doi:10.1080/02640414.2018.1500425.

Whitehead S, Till K, Weaving D, Jones B. 2018. The use of microtechnology to quantify the peak match demands of the football codes: a systematic review. Sports Med (Auckland, NZ). 48(11):2549–2575. doi:10.1007/s40279-018-0965-6.

Paper III

An analysis of training load in highly trained female football players

Andreas K. Winther[1*], Ivan Baptista[2,3], Sigurd Pedersen[1], João Brito[4], Morten B. Randers[5], Dag Johansen[2], Svein Arne Pettersen[1]

[1]School of Sport Sciences, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway.

[2]Department of Computer Science, Faculty of Science and Technology, UiT The Arctic University of Norway, Tromsø, Norway.

[3]Center for Research, Training, Innovation and Intervention in Sport (CIFI2D), Faculty of Sport, University of Porto, Porto, Portugal.

[4]Portugal Football School, Portuguese Football Federation, Lisbon, Portugal.

[5]Department of Sports Science and Clinical Biomechanics, SDU Sport and Health Sciences Cluster, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark.

***Correspondence:**

Andreas K. Winther

andreas.k.winther@uit.no

Keywords: training load, football, soccer, locomotor demands, periodization, performance.

Abstract

This observational study aimed to analyze external training load in highly trained female football players, comparing starters and non-starters across various cycle lengths and training days.

Method: External training load [duration, total distance [TD], high-speed running distance [HSRD], sprint distance [SpD], and acceleration- and deceleration distance [$AccDec_{dist}$] from 100 female football players (22.3 ± 3.7 years of age) in the Norwegian premier division were collected over two seasons using STATSports APEX. This resulted in a final dataset totaling 10498 observations after multiple imputation of missing data. Microcycle length was categorized based on the number of days between matches (2 to 7 days apart), while training days were categorized relative to match day (MD, MD+1, MD+2, MD-5, MD-4, MD-3, MD-2, MD-1). Linear mixed modeling was used to assess differences between days, and starters vs. non-starters.

Results: In longer cycle lengths (5-7 days between matches), the middle of the week (usually MD-4 or MD-3) consistently exhibited the highest external training load (~21-79% of MD TD, MD HSRD, MD SpD, and MD $AccDec_{dist}$); though, with the exception of duration (~108-120% of MD duration), it remained lower than MD. External training load was lowest on MD+2 and MD-1 (~1-37% of MD TD, MD HSRD, MD SpD, MD $AccDec_{dist}$, and ~73-88% of MD peak speed). Non-starters displayed higher loads (~137-400% of starter TD, HSRD, SpD, $AccDec_{dist}$) on MD+2 in cycles with 3 to 7 days between matches, with non-significant differences (~76-116%) on other training days.

Conclusion: Loading patterns resemble a pyramid or skewed pyramid during longer cycle lengths (5-7 days), with higher training loads towards the middle compared to the start and the end of the cycle. Non-starters displayed slightly higher loads on MD+2, with no significant load differentiation from MD-5 onwards.

Introduction

In high-performance sports, a key challenge for coaches and players is striking the right balance between training and recovery. On the one hand, increases in duration, frequency, and intensity of training are often associated with an enhancement in performance [1-3]. On the other hand, increases in training load without adequate recovery may hinder performance and increase the potential risk of injury [4]. In order to find this balance, the periodization of training is considered to be a critical tool [5]. This involves sequencing the overall training plan into units of different lengths (i.e., macro- and microcycles) and planning specific training activities and intensities for each unit [6, 7]. Through careful planning and monitoring, players are believed to maintain a healthy balance between pushing their physical limits and allowing for adequate recovery [5].

In football, variations in training load are most frequently seen at the microcycle level [8]. This is because microcycles can easily be manipulated based on the number of days between matches, allowing practitioners to plan loads that provide a physical stimulus to the players and facilitate recovery [8, 9]. More recently, a principle known as "horizontal alternation" [10] has often been mentioned in tandem. This principle encompasses the idea that physical capacities such as strength, endurance, or speed are targeted on specific days, potentially maximizing the stimuli of each capacity while at the same time minimizing any physiological interferences[11]. This is often done within "days before the match" (MD-) and/or "days after a match" (MD+) framework. To give an example, with six days between matches, three "acquisition" days (MD-4, MD-3, and MD-2) could be placed in between one or two "recovery" days (MD+1 to MD+2) and one "tapering" day (MD-1), where then each "acquisition" day could be dedicated to a specific capacity. In this way, all capacities are maintained or further developed while allowing players enough time to recover between matches.

In professional football, the widespread adoption of Global Positioning System (GPS)-based tracking systems has become prevalent for monitoring the players' activity profiles. These systems can provide practitioners with numerous metrics about a player's external training load and are considered valid and reliable in this respect [12]. Regarding the metrics themselves, both total distance and metrics describing distances covered at various speeds and accelerations and decelerations are typical metrics that both coaches and players want to see [13, 14]. This extends to the planning and monitoring of training, where said metrics could be used as indicators for whether a physical capacity was appropriately targeted. For example, one could expect a more "strength" oriented day to coincide with more accelerations and decelerations [15] due to smaller pitch sizes allowing for more duals and changes of direction. In the same manner, a more endurance-focused day could coincide with longer training durations and total distance covered and a "speed" day with more distance covered at higher speeds [16].

To date, only a few studies have analyzed the periodization of training load in women's football. Most recently, Karlsson et al. [15] found that a Norwegian team differentiated their training load in longer cycles (with 5-7 training days available), closely resembling the horizontal alternation principle. In cycles with four days between matches, Diaz-Seradilla et al. [17] found that MD was more demanding than any training day, while all external training load variables were higher on MD-3 compared to any other training day. Romero-Moraleda et al. [18] also found that the match was the most demanding session in cycles with five days between matches while observing that the training load followed a pyramid shape in which the MD-4 and MD-3 consistently produced the greatest physiological and biomechanical loads, and MD+1 the lowest values.

While research has examined differences between training days in some cycle lengths, little is known about the training load across a broad range of cycles. Furthermore, all previous studies have only investigated players with over 60 minutes of playing time, meaning little is known about the training load of non-starters. Thus, this study aimed to analyze the external training load across a range of typical cycle lengths in professional football, including potential differences between starters and non-starters. We hypothesized that teams differentiated their training load, especially during longer cycles, and that non-starters had higher training loads on MD+1 and MD+2.

Methods

Before commencing the study, we applied for ethical approval through the Regional Committee for Medical and Health Research Ethics - Northern Norway (reference number 53884). We were exempted since the data collection did not include a biobank, medical or health data related to illness, or interfered with the regular operation of the players. After approval from the Norwegian Centre for Research Data (reference number: 296155), we obtained written informed consent from 100 female football players (22.3 ± 3.7 years of age) representing four teams in the Norwegian premier division, classified as highly trained according to the criteria outlined by McKay et al. [19]. Starting in March 2020, a prospective observational study was conducted in which tracking data from training and matches over two full seasons were collected using STATSports Apex (Newry, Northern Ireland), with a sampling frequency of 10 Hz. The validity and level of accuracy (bias <5%) of this tracking system have been previously presented [20]. All teams trained and played home matches on artificial grass, with only occasional away games on natural grass. Training sessions were usually started between 10 AM and 4 PM, with matches typically played between 1 PM and 9 PM. During training and matches, players wore their GPS unit on their upper back, adhering to manufacturer instructions. Furthermore, to minimize inter-device errors [20], each player used the same GPS unit throughout data collection. For the study, we only included outfield players and players with at least one appearance in an official match lineup, either as a starter or as a bench player.

Data pre-processing

Following GPS reporting standards [21], we exported raw GPS data from the manufacturer's software (STATSports Sonra 2.1.4, Newry, Northern Ireland) into a Python (3.9.12) script for pre-processing. Here, we applied a 1-second moving average to smooth doppler-derived speed and derive distance and acceleration. Next, another custom script calculated the physical performance variables. These included duration (measured using timestamps from the raw data), peak speed, total distance (TD), high-speed running distance (HSRD) (>16 km·h^{-1}), and sprint distance (SpD) (>20 km·h^{-1}) based on previous research [22-24]. In addition, combined acceleration- and deceleration distance (AccDec$_{dist}$) was defined as the distance covered with a positive or negative change in speed of more than ± 2.26 m·s^{-2}, finishing when the rate of acceleration/deceleration reached 0 m·s^{-2}.

After deriving all the metrics, the data were transferred to an R 4.0.5 [25] script for missing data imputation and statistical analysis. All variables included in the final analysis are listed in Table 1.

Table 1. Overview of variables included in the final analysis.

Variable	Threshold	Type	Units
Duration		Continuous	Minutes (min)
TD		Continuous	Meters (m)
HSRD	>16 km•h^{-1}	Continuous	m
SpD	>20 km•h^{-1}	Continuous	m
AccDec$_{dist}$,	>2.26 m•s^{-2}	Continuous	m
Peak speed		Continuous	Meters per seconds (m·s^{-1})

Match day and cycle		Nominal	MD, MD+2x3, MD-1x3, MD+2x5, MD-3x5, MD-2x5, MD-1x5, MD+2x6, etc.
Squad status		Nominal	Starter, non-starter
Player ID		Nominal	
Team ID		Nominal	

TD – Total distance; HSRD – high-speed running distance; SpD – Sprint distance; $\text{AccDec}_{\text{dist}}$ – Acceleration and Deceleration distances; MD – Match-day

Handling of missing data

To handle missing data, we followed recommendations by Bache-Mathiesen et al. [26], Borg et al. [27], and Malone et al. [21]. First, we set all physical performance variabbles as missing on sessions with a mean horizontal dilution of precision >5 or a mean number of satellites <12. We also set peak speed as missing if above 32 km·h^{-1} based on theoretical max speed values of 29.2 ± 1.4 km·h^{-1} in a similar cohort [28].

The initial dataset included one observation for each squad player for each day throughout the competitive season (lasting 157 and 176 days in 2020 and 2021, respectively), totaling 12879 observations, with 7646 missing. We opted to remove all observations on MD+1 since it typically was a day off with a substantial amount of missing data (2208 out of 2426 observations). We also removed all observations in cycles with four training days due to too few observations (171 in total with 132 missing). An overview of missing values in the final dataset is shown in Table 2.

Table 2. Number of missing and non-missing observations.

MD (+-)	Cycle	Total non-missing	Total missing	# of players	Mean # of non-missing obs. p/player	Mean HDOP	Mean # of satellites
MD		1158	527	100	11.9	1.3	18.7
MD + 2	2	181	106	95	3.5	2.0	19.8
MD + 2	3	228	413	100	2.9	1.5	19.6
MD - 1		392	247	100	4.8	1.7	19.6
MD + 2	5	133	415	100	1.8	1.7	18.5
MD - 3		333	215	100	3.4	1.6	19.1
MD - 2		208	340	100	2.7	1.6	18.9
MD - 1		314	234	100	3.3	1.8	19.4
MD + 2	6	210	391	99	2.7	1.5	18.7
MD - 4		366	235	99	3.9	1.5	18.5
MD - 3		346	254	99	3.8	1.5	18.6
MD - 2		145	455	99	3.5	1.1	17.7
MD - 1		280	320	99	3.7	1.6	19.5
MD + 2	7	69	243	100	1.6	1.4	17.6
MD - 5		174	138	100	2.0	1.6	18.7
MD - 4		127	185	100	1.7	1.5	18.9
MD - 3		103	209	100	1.7	1.7	19.8
MD - 2		56	256	100	1.1	2.0	18.9
MD - 1		153	159	100	2.1	1.5	19.8

MD – match-day; # - number; HDOP – horizontal dilution of precision; p/player – per player

We used multiple imputation with predicted mean matching (PMM) to impute the missing data, consistent with Bache-Mathiesen et al. [26]. Using the mice package [29] in R, we applied the PMM (mice.impute.pmm) method, including all dependent variables in addition to day number, to generate five imputed datasets for subsequent analysis.

Statistical analysis

Duration, TD, peak speed, and $AccDec_{dist}$ were modelled in R using the lmer package [30], while HSRD and SpD were modelled in the same software using glmmTMB [31]. All models included the interaction between match day, cycle, and squad status as fixed effects and player ID and team ID as random effects. In addition, HSRD and SpD were modelled using the tweedie family with a log link function. Next, we examined, only for the starters, the differences in training load between each day within each cycle and then compared the differences in training load between starters and non-starters within each day. Here, the package emmeans [32] was used to compute estimated marginal means, using the Sidak method to adjust for multiple comparisons between the days and the Tukey method for pairwise comparison between starter and non-starters. We also conducted the same statistical analysis on the non-imputed dataset with only complete cases for sensitivity purposes. Unless otherwise stated, all results are reported as estimated marginal means ± 95% confidence intervals.

Results

Results from the imputed datasets and subsequent models are shown in supplementary tables S1-S3, and Fig 1. Overall, both multiple imputation and complete case analysis gave similar results, and thus only the multiple imputation results are described below. The results for the complete case analysis can be found in supplementary tables S4-S6, and S1 Fig.

Fig 1. External training load by number of days between matches and in proximity to match day (imputed data).

[INSERT FIG 1 HERE]

Match vs. training

Starters displayed significantly higher values (p < 0.001) for TD, HSRD, SpD, $AccDec_{dist}$ and peak speed on MD compared to any other day. MD duration was approximately 88 ± 1 min, shorter (7 ± 4 to 18 ± 4 min, p < 0.001) than training on most acquisition days (MD-5 to -3) in cycles with 5-7 days between matches.

Three days between matches

With three days available (1280 observations), there were no significant differences in duration and TD between MD+2 and MD-1. However, $AccDec_{dist}$, HSRD, SpD and peak speed were slightly higher on MD-1 compared to MD+2, with differences of 108 ± 91 (p = 0.005), 77 ± 38 m (p < 0.001), 21 ± 13 (p < 0.001), and 2.2 ± 1.2 $km \cdot h^{-1}$ (p = 0.01), respectively.

Five days between matches

In cycles with five days between matches (2192 observations), TD, HSRD, SpD, $AccDec_{dist}$ and peak speed were all lower on MD+2 compared to the other training days, except for TD (81 ± 493 m, p = 1.000) and $AccDec_{dist}$ (47 ± 115 m, p = 1.000) on MD-1. Differences in TD and mean peak speed ranged from 2728 ± 434 to 1005 ± 597 m, and from 2.9 ± 1.9 to 5.0 ± 1.5 $km \cdot h^{-1}$, respectively, whilst differences in HSRD and SpD ranged from 82 ± 49 to 356 ± 74 m and from 24 ± 13 to 108 ± 30 m. Differences in $AccDec_{dist}$ ranged from 668 ± 122 τo 251 ± 121 m. All variables were higher (p < 0.001) on MD-3 compared to the other days of the cycle, with the largest differences observed when compared to MD+2 and MD-1, respectively.

Six days between matches

In six-day cycles (3002 observations), all variables were higher on MD-4 to MD-2 compared to MD+2 (p < 0.001). Similarly, both TD (ranging from 1712 ± 430 to 3087 ± 380 m), HSRD (82 ± 42 to 353 ± 111 m), SpD (20 ± 15 to 102 ± 27 m), and $AccDec_{dist}$ (320 ± 137 to 721 ± 110 m) were higher on MD-4 to MD-2 compared to MD-1. However, statistically non-significant differences in peak speed (0.7 ± 0.8 km·h^{-1}, p = 0.158) were found between MD-4 and MD-1. Furthermore, MD-3 had higher duration (11 ± 5 min, p < 0.001) and higher peak speeds (1.5 ± 1.0 km·h^{-1}, p < 0.001) compared to MD-4, and higher TD (1323 ± 370 and 1374 ± 463 m, $p \leq 0.001$), HSRD (245 ± 65 and 188 ± 84 m, p < 0.001), SpD (81 ± 26 and 58 ± 29 m, p < 0.001) and $AccDec_{dist}$ (244 ± 103 and 408 ± 182 m, p < 0.001) compared to both MD-4 and MD-2. The only difference between MD-4 and MD-2 was in $AccDec_{dist}$ (157 ± 156 m, p = 0.047) and peak speed (1.1 ± 1.1 km·h^{-1}, p = 0.047), with higher $AccDec_{dist}$ covered on MD-4, and higher peak speed on MD-2.

Seven days between matches

Seven-day cycles (1872 observations) saw a similar pattern to that of five and six, with all variables being higher on MD-5 to MD-3 compared to MD+2. There were also differences in the tapering stage of the cycle, with longer (11 ± 10 min, $p \leq 0.015$) practice time on MD-2 compared to MD-1, coupled with more TD (1021 ± 602 m, p < 0.001) and $AccDec_{dist}$ (162 ± 142 m, p = 0.009) covered. TD, HSRD, SpD and $AccDec_{dist}$ were higher on MD-4 than any other training day. $AccDec_{dist}$ was higher MD-5 versus MD-3 (211 ± 171 m, p = 0.004).

Starters vs non-starters

Starters vs. non-starters displayed mostly small and non-significant differences in external training load, except on MD+2. Non-starters trained longer (7 ± 5 to 13 ± 4 min, $p \leq$ 0.001) in cycles with 3-6 days between matches, resulting in more TD (731 ± 246 to 1197 ± 218 m, p < 0.001), $AccDec_{dist}$ (176 ± 68 to 346 ± 106 m, p < 0.001), HSRD (28 ± 23 to 51 ± 26 m, $p \leq 0.019$) and higher peak speeds (1.2 ± 1.2 to 1.7 ± 0.7 km·h^{-1}) on those days.

Discussion

Our study aimed to analyze the external training load across a range of typical cycle lengths in professional football, including potential differences between starters and non-starters. We hypothesized that teams differentiated their training load, especially during longer cycles, and that non-starters had higher training loads on MD+1 and MD+2. In line with this, two major findings were apparent from this study. First, the results indicate that the teams in our study altered their external load based on the number of days between matches, with most of the training load clustered towards the mid-week, succeeding and preceding days of lower loads. Secondly, there was little to no differentiation in training load between starters and non-starters from MD-5 and onwards, regardless of cycle.

Our data indicates that teams perform the highest combined external load at least three to four days pre-match in a typical match fixture. This period of higher load succeeds and precedes days of lower load, which makes sense from a periodization standpoint. This forms a basic structure where the first few days post-match are usually geared towards recovery, mid-week towards acquisition, and pre-match towards tapering. However, in shorter cycles with only two or three days between matches, our data indicates that most of the time is spent at lower loads awaiting a mid-week game. Regardless of metric, however, the loads are lower than match day, though there is less difference for $AccDec_{dist}$ than for SpD. For example, $AccDec_{dist}$ on MD-3 and MD-4 in longer cycles (5-7 days between matches) is 59-79% of MD, while SpD is only 22-51%. This could be due to a preference for small-sided games, which often involve a smaller area, hence giving insufficient space to accumulate distances at higher speeds [33].

Regarding the training day differentiation, we found no apparent differences in SpD between MD-4 and MD-2 and between MD-5 and MD-3 in cycles of six and seven days between matches. However, for $AccDec_{dist}$, we did find significantly more distance covered on MD-4 and MD-5 versus MD-2 and MD-3, which could suggest a day with smaller spaces, while there was a tendency for higher peak speeds on MD-2 and MD-3. It is also interesting to note that the highest estimated mean peak speed in training was 24.2 km·h^{-1} or ~93% of the estimated mean peak speed on MD for starters (26.0 km·h^{-1}). Considering that Haugen et al. [28] found theoretical peak speed values of 29.2 ± 1.4 km·h^{-1}, or roughly ~112% of match day, we are looking at a difference of ~19% between what players could be physically able to achieve versus what they are achieving in training. Added to the fact that training load decreases throughout the season [15, 34], this could explain why we also see a concurrent decrease in sprint ability during this period [34]. From a specificity standpoint, coaches should be aware of the importance of training at maximum running speed to enhance this capacity [35]. In addition, being exposed to maximum speeds could also be important from an injury prevention standpoint, as exposure to high-speed football actions has been suggested to be a modifiable risk factor for hamstring injuries [36, 37].

Continuing with the second finding, it is more challenging to discern whether load compensation is given for the substitutes in the combination of cycles and days available. Although there were some differences between starters and substitutes in training duration, TD, peak speed, and $AccDec_{dist}$ on MD+2 in most cycles, this could be due to residual fatigue from the last match in the starters. There were also no pronounced differences in HSRD and SpD, and the overall load was considerably than any other day. However, this does not exclude the fact that substitute compensation could have occurred at MD or MD+1 or both. For example, training could be executed in forms that do not require tracking equipment since there were huge amounts of missing data at MD+1. Of note is that the teams in our sample had both a second team and a junior team, and it is likely that match play at these competitive levels was given as compensation, again, without it being tracked.

Our results are comparable to other studies on female football players, most notably to Karlsson et al. [15], which also included a team from the Norwegian premier division. Overall, the loading patterns and distances covered were fairly similar, with MD-4 and MD-3 dependent on cycle, equivalent to MD-3 in Karlsson et al. [15], being the training day with the highest external load for most metrics. That external training load is higher on MD-3 in cycles with five days between matches is also consistent with Diaz-Seradilla et al. [17]. Karlsson et al. [15] also found a higher number of accelerations and decelerations on the day preceding and more sprint distance covered on the day succeeding the highest overall day, which we did not observe considering SpD covered, although metrics are not directly comparable. That MD contains the highest external load is also supported by previous studies [15, 17, 18]. Compared to studies on male players, however, our results are similar to Akenhead et al. [38] and Anderson et al. [39], who examined the external training load of English Premier League teams. Together, they both show a pattern of higher loads preceding and succeeding days of lower loads irrespective of metric, similar to our study. In addition, in a study on a team from the Eredivisie, Stevens et al. [40] noted that relative to match values (100%), accelerations and decelerations (39-90%) were much higher compared to the other metrics, which mirrors our study (27-79%).

A major strength of this study is that we utilized a multi-team, multi-season approach, in contrast to most other observational studies in football, which usually are one-team, one-season. We also examined external training load across a range of cycle lengths with different numbers of days between matches. This complements Karlsson et al. [15], who concatenated similar days in cycles with five, six, and seven days between matches while adding to Diaz-Seradilla et al. [17] and Romero-Moraleda et al. [18], who studied cycles with four and six days available, respectively. Another strength is that we compared external training load for both starters and non-starters, which, as far as we know, has not been investigated in women's elite football.

There are also some limitations to our approach. Mainly, we lacked context surrounding each training day, thus making it very hard to discern whether training had occurred or not on observations with missing data. Therefore, some observations were likely imputed when they should have been removed from the dataset. Also, the categories of starters and non-starters could be viewed as somewhat crude. For example, we put starters who were subbed out early and those who played the whole game within the starter category. However, if we were to dichotomize further, this would only run into the problem of where to set the cut-off point, and thus, we thought it was better to leave this category as it was.

Our study serves as a springboard for future research endeavors to refine our understanding of external training load dynamics in professional football. To enhance precision, future studies should gather contextual information surrounding each training day, including details on specific drills and focus areas. Refining player categories to capture more nuanced distinctions, such as players substituted early versus those playing the whole game, can provide additional insights into training load variations. In addition, longitudinal studies spanning multiple seasons and teams can reveal evolving trends influenced by changes in coaching staff, coaching strategies, or other external factors. Finally, investigating the impact of different training formats, including small-sided games and specific drills, on external training load can guide coaches in designing more effective training sessions.

Practical applications

The insights gained from our study have several practical implications for the training and management of professional football players. First, our study identifies a sound approach to training load periodization, wherein the emphasis of a high load day, succeeding and preceding days of lower load, is well within the recommendations of contemporary training theory. Practitioners can thus leverage the findings of our study as a template or starting point when designing their training programs. Furthermore, our observation of a possible disparity between the players' maximum achievable speed and their training speeds highlights an area for improvement. Combined with the fact that the volume of sprinting is comparably low especially regarding acceleration, this could point to a neglect of speed work in the daily training regimen of female football players. Thus, coaches should be mindful of incorporating exercises that allow players to reach top speeds. This not only enhances speed-specific capacities but may also contribute to injury prevention. Finally, our findings raise awareness of potential load compensation for non-starters, prompting coaches to explore compensatory strategies for optimal player development.

Conclusion

These results provide further evidence regarding how highly trained female football teams adjust their external training load across various microcycles. Loading patterns typically take on a shape similar to a pyramid, or a skewed pyramid, during longer cycle lengths, with higher training loads towards the middle compared to the start and the end of the cycle. Non-starters reached higher peak speeds and covered more total distance and combined acceleration- and deceleration distance on MD+2 in cycles with 3-6 days between matches. However, there was no significant load differentiation from MD-5 and onwards.

Acknowledgement

We would like to thank Luciano Arias, Helene Davidsen Moa, Astrid G. Ree, Elias Koppen, Marte Vatn, Thomas Lyngbø, and Atle Ingebrigtsen for their assistance in data collection and the players in their respective clubs for approving to participate in the study.

References

1. Pollock ML, Gettman LR, Milesis CA, Bah MD, Durstine L, Johnson RB. Effects of frequency and duration of training on attrition and incidence of injury. Med Sci Sports Exerc. 1977;9(1):31-6.
2. Mujika I. Intense training: the key to optimal performance before and during the taper. Scand J Med Sci Sports. 2010;20:24-31.
3. Branquinho L, Ferraz R, Marques MC. 5-a-Side Game as a Tool for the Coach in Soccer Training. Strength Cond J. 2021;43(5):96-108.
4. Meeusen R, Duclos M, Foster C, Fry A, Gleeson M, Nieman D, et al. Prevention, diagnosis and treatment of the overtraining syndrome: Joint consensus statement of the European College of Sport Science (ECSS) and the American College of Sports Medicine (ACSM). Eur J Sports Sci. 2013;13(1):1-24.
5. Kiely J. Periodization paradigms in the 21st century: evidence-led or tradition-driven? Int J Sports Physiol Perform. 2012;7(3):242-50.
6. Mujika I, Halson S, Burke LM, Balagué G, Farrow D. An integrated, multifactorial approach to periodization for optimal performance in individual and team sports. Int J Sports Physiol Perform. 2018;13(5):538-61.
7. Issurin VB. New horizons for the methodology and physiology of training periodization. Sports Med. 2010;40:189-206.
8. Clemente FM, Martins FM, Mendes RS. Periodization based on small-sided soccer games: Theoretical considerations. Strength Cond J. 2014;36(5):34-43.
9. Morgans R, Orme P, Anderson L, Drust B. Principles and practices of training for soccer. J Sport Health Sci. 2014;3(4):251-7.

10. Buchheit MSM, Berndsen J, Shelton SS, Norman D, McHugh D, Hader K. Loading patterns and programming practices in elite football: insights from 100 elite practitioners. Sport Perf Sci Rep. 2021;153:1-18.

11. Fyfe JJ, Bishop DJ, Stepto NK. Interference between concurrent resistance and endurance exercise: molecular bases and the role of individual training variables. Sports Med. 2014;44:743-62.

12. 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):1470-90.

13. Akenhead R, Nassis GP. Training load and player monitoring in high-level football: current practice and perceptions. Int J Sports Physiol Perform. 2016;11(5):587-93.

14. Nosek P, Brownlee TE, Drust B, Andrew M. Feedback of GPS training data within professional English soccer: a comparison of decision making and perceptions between coaches, players and performance staff. Sci Med Football. 2021;5(1):35-47.

15. Karlsson UB, Vagle M, Wiig H, Luteberget LS. Training Load Quantification in Women's Elite Football: A Season-Long Prospective Cohort Study. Int J Sports Physiol Perform. 2023;18(6):603-14.

16. Buchheit M, Lacome M, Cholley Y, Simpson BM. Neuromuscular responses to conditioned soccer sessions assessed via GPS-embedded accelerometers: insights into tactical periodization. Int J Sports Physiol Perform. 2018;13(5):577-83.

17. Diaz-Seradilla E, Rodríguez-Fernández A, Rodríguez-Marroyo JA, Castillo D, Raya-González J, Villa Vicente JG. Inter-and intra-microcycle external load analysis in female professional soccer players: A playing position approach. PLoS ONE. 2022;17(3):e0264908.

18. Romero-Moraleda B, Nedergaard NJ, Morencos E, Casamichana D, Ramirez-Campillo R, Vanrenterghem J. External and internal loads during the competitive season in

professional female soccer players according to their playing position: differences between training and competition. Res Sports Med. 2021;29(5):449-61.

19. McKay AK, Stellingwerff T, Smith ES, Martin DT, Mujika I, Goosey-Tolfrey VL, et al. Defining Training and Performance Caliber: A Participant Classification Framework. Int J Sports Physiol Perform. 2022;17(2):317-31.

20. Beato M, Coratella G, Stiff A, Iacono AD. The Validity and Between-Unit Variability of GNSS Units (STATSports Apex 10 and 18 Hz) for Measuring Distance and Peak Speed in Team Sports. Front Physiol. 2018;9:1288.

21. Malone JJ, Lovell R, Varley MC, Coutts AJ. Unpacking the black box: applications and considerations for using GPS devices in sport. Int J Sports Physiol Perform. 2017;12(S2):S218-S26.

22. Bradley PS, Vescovi JD. Velocity thresholds for women's soccer matches: Sex specificity dictates high-speed-running and sprinting thresholds—Female Athletes in Motion (FAiM). Int J Sports Physiol Perform. 2015;10(1):112-6.

23. Vescovi JD. Sprint speed characteristics of high-level American female soccer players: Female Athletes in Motion (FAiM) study. J Sci Med Sports. 2012;15(5):474-8.

24. Winther AK, Baptista I, Pedersen S, Randers MB, Johansen D, Krustrup P, et al. Position specific physical performance and running intensity fluctuations in elite women's football. Scand J Med Sci Sports. 2022;32:105-14.

25. R Core Team R. R: A language and environment for statistical computing. 2013.

26. Bache-Mathiesen L, Andersen TE, Clarsen B, Fagerland MW. Handling and reporting missing data in training load and injury risk research. Sci Med Football. 2022;6(4):452-64.

27. Borg DN, Nguyen R, Tierney NJ. Missing data: current practice in football research and recommendations for improvement. Sci Med Football. 2022;6(2):262-7.

28. Haugen TA, Breitschädel F, Seiler S. Sprint mechanical properties in soccer players according to playing standard, position, age and sex. J Sports Sci. 2020;38(9):1070-6.

29. Van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. J Stat Soft. 2011;45:1-67.

30. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Soft. 2015;67(1):1-48.

31. Magnusson A, Skaug H, Nielsen A, Berg C, Kristensen K, Maechler M, et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R J. 2017;9(2):378-400.

32. Lenth R, Singmann H, Love J, Buerkner P, Herve M. Emmeans: Estimated marginal means, aka least-squares means. R package version 1.5. 2018.

33. Beato M, Vicens-Bordas J, Peña J, Costin AJ. Training load comparison between small, medium, and large-sided games in professional football. Front Spo Act Liv. 2023;5:1165242.

34. Mara JK, Thompson KG, Pumpa KL, Ball NB. Periodization and physical performance in elite female soccer players. Int J Sports Physiol Perform. 2015;10(5):664-9.

35. Young WB, McDowell MH, Scarlett BJ. Specificity of sprint and agility training methods. J Strength Cond Res. 2001;15(3):315-9.

36. Edouard P, Mendiguchia J, Guex K, Lahti J, Samozino P, Morin J-B. Sprinting: a potential vaccine for hamstring injury. Sport Perf Sci Rep. 2019;1:1-2.

37. Ekstrand J, Hallén A, Marin V, Gauffin H. Most modifiable risk factors for hamstring muscle injury in women's elite football are extrinsic and associated with the club, the team, and the coaching staff and not the players themselves: the UEFA Women's Elite Club Injury Study. Knee Surg Sport Trauma Arthro. 2023;31:2550-5.

38. Akenhead R, Harley JA, Tweddle SP. Examining the external training load of an English Premier League football team with special reference to acceleration. J Strength Cond Res. 2016;30(9):2424-32.

39. Anderson L, Orme P, Di Michele R, Close GL, Morgans R, Drust B, et al. Quantification of training load during one-, two-and three-game week schedules in professional soccer players from the English Premier League: implications for carbohydrate periodisation. J Sports Sci. 2016;34(13):1250-9.

40. Stevens TG, de Ruiter CJ, Twisk JW, Savelsbergh GJ, Beek PJ. Quantification of in-season training load relative to match load in professional Dutch Eredivisie football players. Sci Med Football. 2017;1(2):117-25.

Supporting information

S1 Table. Between-day contrasts from imputed data.

S2 Table. Estimated marginal means by MD, cycle, and squad status, from imputed data.

S3 Table. Starter vs. non-starter contrasts from imputed data.

S4 Table. Between-day contrasts from non-imputed data.

S5 Table. Estimated marginal means by MD, cycle, and squad status, from non-imputed data.

S6 Table. Starter vs. non-starter contrasts from non-imputed data.

S1 Fig. External training load by number of days between matches and in proximity to match day (non-imputed data).