CORSIE, M. and SWINTON, P.A. 2023. Reliability of spatial-temporal metrics used to assess collective behaviours in football: an in-silico experiment. Science and medicine in football [online], 7(3), pages 297-305. Available from: https://doi.org/10.1080/24733938.2022.2100460

Reliability of spatial-temporal metrics used to assess collective behaviours in football: an in-silico experiment.

CORSIE, M. and SWINTON, P.A.

2023

Science and Medicine in Football

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

Reliability of spatial-temporal metrics used to assess collective behaviours in football: an in-silico experiment

Martin Corsie & Paul Alan Swinton

To cite this article: Martin Corsie & Paul Alan Swinton (2023) Reliability of spatial-temporal metrics used to assess collective behaviours in football: an in-silico experiment, Science and Medicine in Football, 7:3, 297-305, DOI: 10.1080/24733938.2022.2100460

To link to this article: https://doi.org/10.1080/24733938.2022.2100460

Published online: 17 Jul 2022.

SCIENCE AND MEDICINE IN FOOTBALL
2023, VOL. 7, NO. 3, 297–305
https://doi.org/10.1080/24733938.2022.2100460

Routledge
Taylor & Francis Group

ORIGINAL INVESTIGATION

OPEN ACCESS Check for updates

Reliability of spatial-temporal metrics used to assess collective behaviours in football: an in-silico experiment

Martin Corsie and Paul Alan Swinton

School of Health Sciences, Robert Gordon University, Aberdeen, UK

ABSTRACT

Background: The purpose of this study was to investigate the reliability of spatio-temporal measurements applied within collective behaviour research in football.

Methods: In silico experiments were conducted introducing positional errors (0.5, 2 and 4 m) representative of commercial tracking systems to match data from the 2020 European Championship qualifiers. Ratios of the natural variance ('signal') of spatio-temporal metrics obtained throughout sections of each game relative to the variance created by positional errors ('noise') were taken to calculate reliability. The effects of error magnitude and time of analysis (1, 5 and 15 mins; length of attack: <10, 10–20, >20 s) were assessed and compared using Cohen's f^2 effect size.

Results: Error magnitude was found to exert greater influence on reliability (f^2 = 0.15 to 0.81) compared with both standard time of analysis (f^2 = 0.03 to 0.08) and length of attacks (f^2 = 0.15 to 0.32).

Discussion: The results demonstrate that technologies generating positional errors of 0.5 m or less should be expected to produce spatio-temporal metrics with high reliability. However, technologies that generate errors of 2 m or greater may produce unreliable values, particularly when analyses are conducted over discrete events such as attacks, which although critical, are often short in duration.

ARTICLE HISTORY
Accepted 6 July 2022

KEYWORDS
Soccer; simulation; position tracking; dynamic system theory

Introduction

Player tracking has become a staple of performance analysis in elite team-based sports(Buchheit and Simpson 2017) and has traditionally been used to help coaches understand the physical demands and activity profiles of players(Cummins et al. 2013; Sarmento et al. 2018). However, a focus on individual and physical outputs provides limited information regarding the overall functioning of the team. Consequently, collective behaviour metrics quantifying team behaviour and decision-making have emerged as an alternative application of player tracking data. In team sports, this is commonly achieved by Global Positioning Systems (GPS), semi-automatic video tracking and radio-based local position measurement systems. The information obtained provides players' position in space at relatively high frequencies (e.g., 1 to 45 Hz) and hence is referred to as spatio-temporal data and has the potential to describe collective behaviour (Low et al. 2020). A range of sports have been explored (Sampaio and MaçÃs 2012; Gonçalves et al. 2016, 2019; Moura et al. 2016; Laakso et al. 2017); however, one of the most popular to apply collective behaviour metrics is football, where the complexity of the game and relatively low scoring opportunities has limited the success of traditional performance analysis methods such as frequency analysis in understanding team behaviour (Sarmento et al. 2018). A range of potential applications have been proposed for spatio-temporal data including talent identification (Low et al. 2020), evaluation of complex decisions made by players (Steiner et al. 2019), development of youth players (Barnabé et al. 2016), enhanced training sessions through constraint manipulation (Silva et al. 2014b), and even providing live data to inform coach decision making during matches (Clemente et al. 2013). However, extensive theoretical and practical development is required before approaches are widely used within football teams.

Only recently has research investigating spatio-temporal data and associated team collective behaviours in football been systematically reviewed. Low et al. (2020) created a taxonomy with 6 categories to delineate the most common spatio-temporal metrics reported. Four of these categories considered tactical variables describing characteristics of player and team organisation. These categories included position, distance, space, and numerical relations that referred to coordinate location, distances between positions, team dispersion and areas of superiority or inferiority, respectively. The remaining two categories (synchronisation and predictability) identified by Low et al. (2020) represent non-linear methods used to analyse metrics from the first 4 categories. Low et al. (2020) identified 27 distinct tactical metrics presented in research investigating collective behaviours in football with spatio-temporal data. However, this categorisation might fail to account for subtle differences between metric calculations that may ultimately influence findings and practical applications.

In addition to the large number of spatio-temporal metrics and analysis methods that can be applied, previous research also varies substantially in the approaches used to process and report analyses including team structure (e.g., dyads, subgroups, the team or across both teams), different periods of competition (e.g., every 5 or 10 minutes, per half, or full game),

CONTACT Paul Alan Swinton p.swinton@rgu.ac.uk School of Health Sciences, Robert Gordon University, Aberdeen, UK

or according to different game scenarios (e.g., all play, continuous play, attacks, or successful attacks). Generally, researchers have recorded metrics at high frequencies and summarised values through means and standard deviations over a range of time intervals including 30 seconds (Clemente et al. 2018), 5 minutes (Moura et al. 2013, 2016), 10 minutes (Moura et al. 2012) and 15 minutes (Duarte et al. 2013; Gonçalves et al. 2018). Statistical analyses are then frequently performed by comparing samples across independent variables that may indicate expertise such as age (Olthof et al. 2015; Barnabé et al. 2016; Aquino et al. 2016a; Menuchi et al. 2018). The most common metrics analysed using this approach include those describing behaviour at the team level such as team centroid (Frencken et al. 2011; Moura et al. 2012; Barnabé et al. 2016; Castellano et al. 2016; Praça et al. 2016; Olthof et al. 2019), surface area (Baptista et al., 2020; Frencken et al. 2011, 2013; Travassos et al. 2014; Aquino et al. 2016b; Castellano et al. 2017) and stretch index (Bartlett et al. 2012; Frias 2014; Silva et al. 2014a; Olthof et al. 2019). Alternative approaches have included calculation of samples across discrete sequences within games such as attacks. This approach has often featured in small-sided games (SSGs) or sub-phases of full games to analyse 1v1 or 2v1 situations with metrics calculated at the dyad level (Headrick et al. 2012; Leser et al. 2015; Laakso et al. 2019). Data processing strategies restricting analyses to the longest attacks (Chung et al. 2019) or successful attacks (e.g., possession resulting in a shot or goal) (Bartlett et al. 2012; Moura et al. 2016; Shafizadeh et al. 2016) have also been used in attempts to assess behaviours over patterns of play that are deemed most meaningful. Clear processes and guidelines have yet to be established due to the developing nature of the research topic.

An important preliminary concern that is yet to be addressed with analysis of spatiol-temporal metrics is reliability given errors in positional data. Where reliability has been discussed previously, this has been restricted to errors in individual players' x-y position that has been shown to range from 2 (Frencken et al. 2010) to 470 cm (Siegle et al. 2013). However, the extent to which these individual position errors influence reliability of specific collective behaviour metrics is unknown. Moreover, it is important to determine the effects of data processing approaches such as analysis over different time periods versus analysis across specific game scenarios on reliability. Given the intrinsic nature of variability in dynamic systems such as collective behaviours in football, general conceptions of reliability as the ratio of variances representing signal relative to noise (Hämmerer et al. 2013; Lacome et al. 2019; Ryan et al. 2020) are required to establish the influence of positional errors. For common approaches to be appropriate, the variance induced by positional errors should be substantively lower than variance exhibited by metrics across the analysis framework selected (e.g., time or specific game scenarios). An efficient and powerful method to investigate reliability is through in silico experiments where positional errors can be systematically introduced and manipulated to reflect expected errors based on current technology. Variance caused by this manipulation can be compared to variance across different frameworks and the reliability quantified. Therefore, the purpose of this study was to provide a novel reliability assessment of common spatio-temporal metrics used in football and assess the potential moderating effects of procedures such as time periods and game scenarios selected to perform calculations.

Methods

Data collection

Data were collected from seven competitive international matches during the UEFA European championships 2020 qualifying stage. The data comprised player x-y position information gathered at 25 Hz using the TRACAB optical tracking system (Chyronhego, New York). Player coordinates from both teams were recorded longitudinally from −5250 to 5250 cm and laterally from −3400 to 3400 cm and were subsequently transformed into metres. Additionally, time-stamped event data were tagged separately and integrated to identify how the attack started and ended and which team was in possession.

Study protocols

Eight frequently investigated spatio-temporal metrics were selected, comprising two metrics per category identified by Low et al. (2020). The metrics included centroid$_x$ and centroid$_y$ (position), distance between team centroids and team length (distance), surface area and team spread (spaces), length per width ratio and space control (numerical relations). Analyses were performed on samples generated across predetermined time intervals (1, 5, or 15 minutes) and across possessions separated by the attacking team. Three criteria were used to identify the start of possessions: (i) where a player takes three consecutive touches of the ball, (ii) a player makes a successful pass to a teammate, or (iii) a player takes a shot (Santos and Lago-Penas 2019). A possession lasted until the ball went out of play or the opposition recovered the ball.

Centroid coordinates were calculated using the following equation:

$$Centroid_{x,y}(t) = \left(\frac{\sum_i^N p_{xi}(t)}{N}, \frac{\sum_i^N p_{yi}(t)}{N} \right). \quad (1)$$

where N is the number of outfield players in the team and $p_{xi}(t), p_{yi}(t)$ represent the x-y position of player i at time t (t). The radial distance between centroids was calculated as

$$\text{Distance between centroids}(t) = \sqrt{(xC(t)_a - xC(t)_b)^2 + (yC(t)_a - yC(t)_b)^2}. \quad (2)$$

where $C(t)_a$ and $C(t)_b$ refer to the centroid coordinate of team a and b, respectively. The team length was calculated as the maximum distance along the x-axis between two outfield players in a team.

Surface area was calculated by creating a convex hull around the outfield players of each team and calculating the area inside (Palucci Vieira et al. 2019). Team spread was measured by calculating the Euclidean distances between all outfield players, arranging the values in a 10 by 10 matrix. The Frobenius norm was calculated using the lower triangular matrix as shown in the following equation :

$$\|\ \mathrm{Team\,spread}(t)\|_F = \sqrt{\sum\nolimits_{i=1}^{N}\sum\nolimits_{j=1}^{N}|l_{ij}|^2} \quad (3)$$

where l_{ij} is the distance between each pair of teammates and N is the number of outfield players (Moura et al. 2016). The length per width ratio (LPW) was calculated as the ratio of team length (maximum distance between teammates in the x direction) relative to the team width (maximum distance between teammates in the y direction).

Space control was measured using Voronoi cell computations that identify the area on the pitch closest to each individual. This describes a player's dominant region, where space control equals the ratio of the sum of team a's dominant regions relative to the sum of team b's dominant regions.

A set of pilot analyses were conducted comparing the reliability values obtained when using data collected at 1, 15 and 25 Hz on reduced sample sizes. Trivial differences in reliability values were obtained such that full analyses were conducted with data reduced to 1 Hz to lower computing requirements. These data were used to calculate the 'natural variance' in metrics (e.g., signal) reflecting the amount of change across time. To complete the in silico experiments, positional errors were added to each players' x-y position at each time recording (e.g., noise). Positional errors comprised random draws from a Gaussian distribution with a mean of 0 and standard deviation selected to represent the typical error magnitude of positional tracking equipment. Three error magnitudes were selected (small, medium, and large) equal to 0.5, 2 and 4 m (Frencken et al. 2010; Ogris et al. 2012; Siegle et al. 2013; Linke et al. 2018, 2020). The process of adding positional errors to all players positional data was completed 1,000 times, such that metrics were calculated on the actual game data and across 1,000 in silico experiments. A schematic overview of the entire in silico approach is illustrated in Figure 1. Reliability values were calculated at each instance of the analysis framework selected (e.g., time period or game scenario) by comparing the variance in the metric across time from *within* the raw data (e.g., signal), relative to the variance in the metric *between* the in silico experiments (e.g., noise). Reliability was calculated as the within variance relative to the total variance (within + between) providing a value between 0 and 1. Values higher than 0.909 were deemed highly reliable (Beltrame et al. 2020).

Statistical analysis

For each metric, the effects of error magnitude (0.5, 2 and 4 m) and time of analysis (playing time: 1, 5 and 15 min; attack duration: short [<10 s], medium [10–20 s], long [≥20 s]) were evaluated by fitting mixed effects beta-regression models with reliability values as the dependent variable and random effects fitted to each game to account for covariance of values made within the same match. Analyses were conducted using the glmmTMB package in R (Brooks et al. 2017). Cohens f^2 effect sizes for mixed effects models were used to quantify and compare the influence of each independent variable (Nakagawa and Schielzeth 2012), with threshold values of 0.02, 0.15 and 0.35 used to categorise effects as small, medium and large, respectively. To quantify uncertainty in estimates, bootstrapping at the match level was conducted with 95% confidence intervals calculated from the respective bootstrap sample (n = 1000) quantiles. The structure in the data was

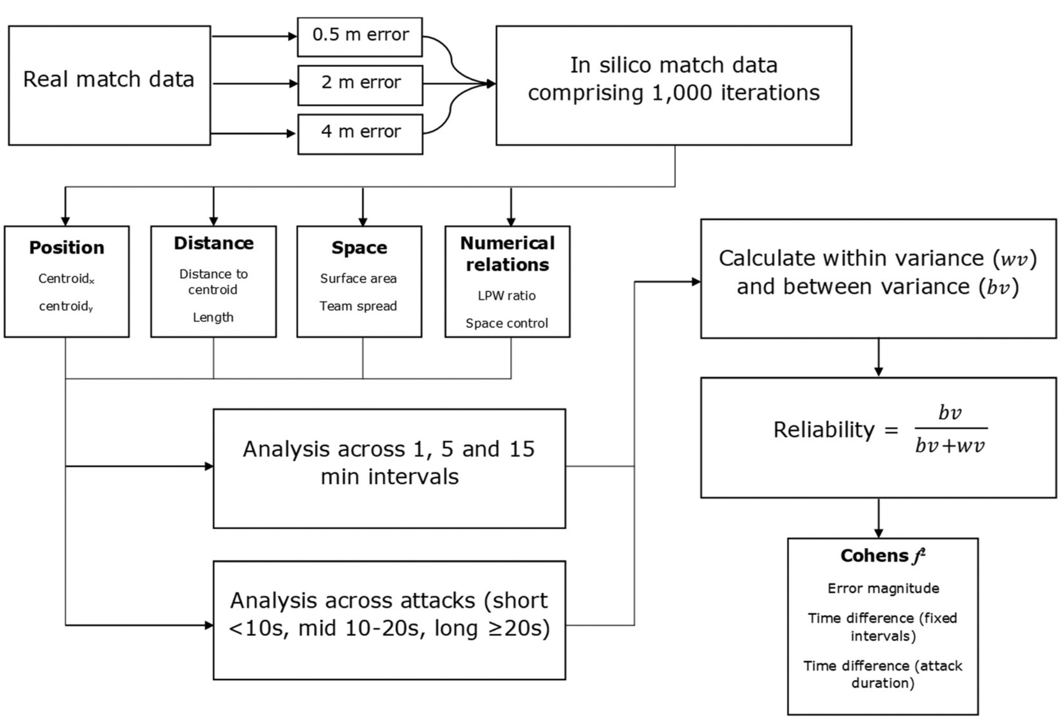

Figure 1. Schematic of the in silico experiment process.

Table 1. Mean (standard deviation) reliability values pooled across metrics and presented by error magnitude and time grouping.

error	1 minute	5 minutes	15 minutes
0.5	0.984 (±0.027)	0.992 (±0.012)	0.992 (±0.012)
2	0.865 (±0.136)	0.924 (±0.077)	0.936 (±0.062)
4	0.706 (±0.215)	0.811 (±0.155)	0.836 (±0.13)

investigated through calculation of variance partition coefficients (VPCs) that compared the variance associated with the random intercept from each game relative to the residual variance. All VPCs, however, were below 0.002 indicating minimal structure related to the individual games.

Results

An initial overview of the results was obtained by aggregating reliability values across all eight metrics and presenting the information relative to the error magnitude and time of analysis (Table 1). A visual distribution of reliability values for individual metrics are presented for different error magnitudes in Figure 2. High mean reliability values (>0.909) were obtained for all metrics with 0.5 m errors. Across all error magnitudes, three clusters were observed with the greatest mean reliability obtained for centroids (x direction: 0.984 to >0.999; y direction: 0.961 to >0.999). In the second cluster, high mean reliability values were obtained with the space metrics (team spread and surface area) for both 0.5 and 2 m errors. In contrast, poor mean reliability was obtained for all other metrics in the third cluster, with values as low as 0.6 for distances between centroids under 4 m errors.

When quantifying the effects of error magnitude and time of analysis on reliability values (Table 2), the relative effects of error magnitude were substantively greater for all metrics. Large effects of error magnitude were identified for all variables (f^2 = 0.56 to 0.81) except for centroids where medium effects were identified (f^2 = 0.15 to 0.28). In contrast, time was found to exhibit small effects on reliability values for all metrics (f^2 = 0.03 to 0.08).

Over the seven games 1189 attacks were analysed, 152 (13%) of which ended in shots, with 21 goals scored. A total of 514 (43%) attacks were categorised as short (<10 s), 294 (25%) were categorised as medium (10–20 s), and 381 (32%) were categorised as long (>20 s). As an initial analysis, reliability values were pooled across all metrics and the effects of error magnitude and duration of attack were visualised (Figure 3). The analysis illustrates that as attack durations decrease, the distribution of reliability values spreads substantively for even 0.5 m errors. In general, reliability values are poor for all conditions except medium- and long-duration attacks with 0.5 m errors.

A similar pattern of results was obtained when comparing the effects of error magnitude and time of analysis on reliability values when restricted to attacks (Table 3). The relative effects of error magnitude were substantively greater for all metrics except centroids where large effects (f^2 = 0.71 to 0.96) were obtained for the

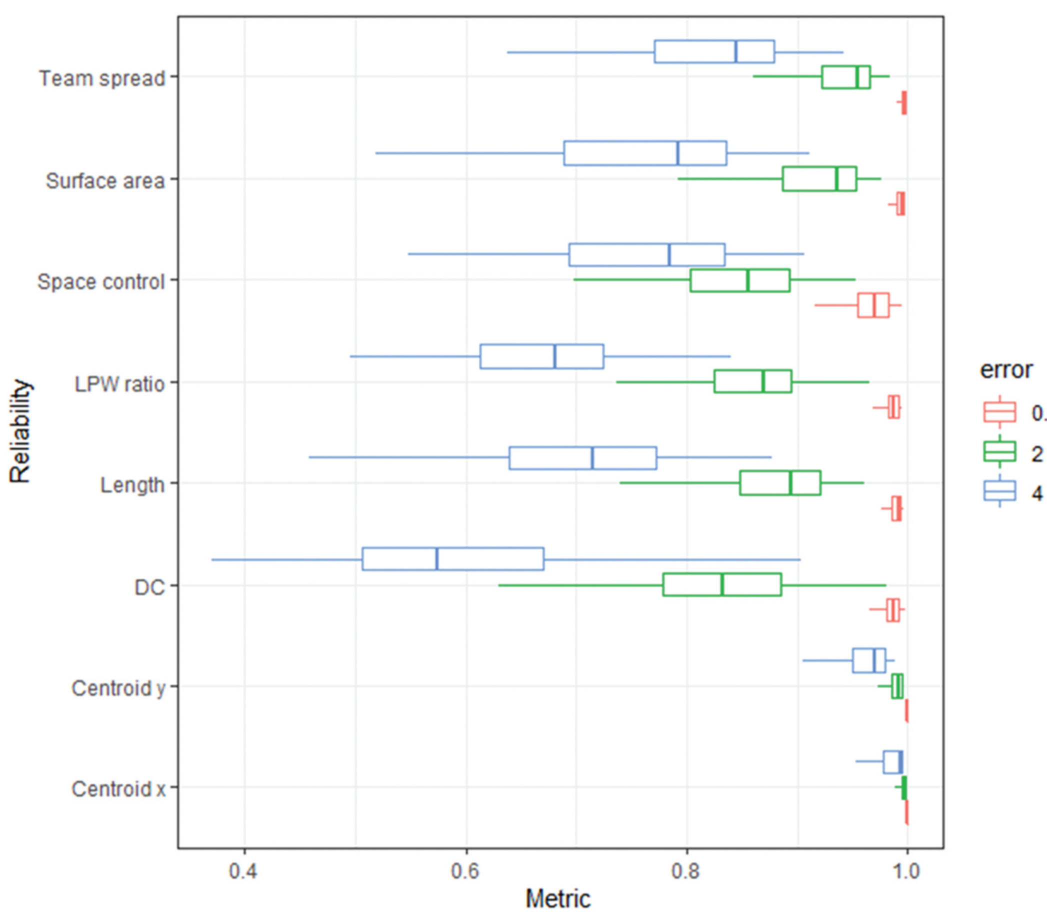

Figure 2. Distribution of reliability values from in silico experiments created with error magnitudes of 0.5, 2 and 4 metres.

Table 2. Effect sizes of error magnitude and time grouping (1, 5 and 15 minutes) on reliability values.

Metric	Time f^2 [95%CI]	Error magnitude f^2 [95%CI]
DC	0.08 [0.05–0.14]	0.77 [0.75–0.81]
Length	0.05 [0.02–0.08]	0.75 [0.73–0.81]
LPW ratio	0.04 [0.02–0.08]	0.81 [0.79–0.83]
Surface area	0.05 [0.03–0.08]	0.64 [0.63–0.66]
Team spread	0.04 [0.02–0.06]	0.56 [0.55–0.58]
Space control	0.06 [0.03–0.09]	0.67 [0.66–0.70]
$Centroid_x$	0.04 [0.03–0.05]	0.15 [0.14–0.16]
$Centroid_y$	0.03 [0.02–0.04]	0.28 [0.27–0.30]

DC = distance between centroid. LPW ratio = length per width ratio.

former and small to medium effects for the latter ($f^2 = 0.10$ to 0.17). In contrast, medium effects of attack duration were obtained for all metrics ($f^2 = 0.15$ to 0.32)

Discussion

The purpose of this study was to investigate the reliability of spatio-temporal metrics used to quantify collective team behaviour in football. Reliability was assessed using in silico experiments with errors representative of that induced by current positional tracking technologies. Collectively, the results demonstrate that technologies, which generate errors in position tracking of 0.5 m or less, should be expected to produce spatio-temporal metrics with high reliability. However, technologies that generate error values of around 2 m may have substantive effects on metrics such as distance between centroids, length, length per width ratio and space control, reducing reliability and limiting the accuracy of conclusions drawn with limited data. Similarly, technologies generating larger positional errors such as 4 m can lower reliability to where noise and signal are approximately equal, severely limiting insights. Whilst researchers and practitioners may seek to analyse spatio-temporal metrics over crucial sequences of plays such as attacks, the findings from this study indicate that this strategy may result in very low reliability, due to the short duration of most attacks. Unacceptably low reliability is likely to occur where analyses are conducted over sequences of plays such as attacks using tracking technologies that generate moderate to large positional errors (e.g., 2 to 4 m).

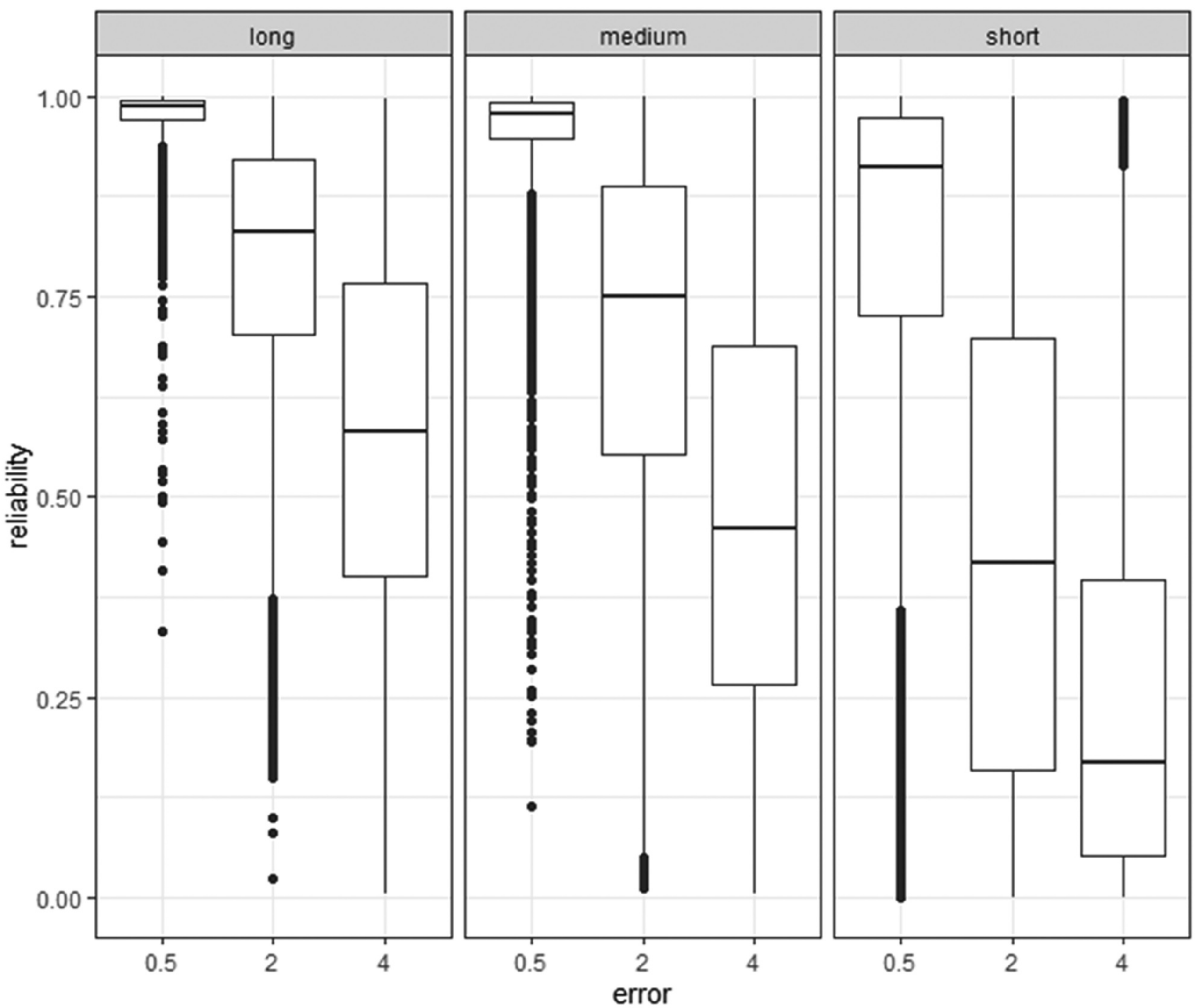

Figure 3. Distribution of reliability values pooled across all metrics from in silico experiments created with error magnitudes of 0.5, 2 and 4 metres and attack time grouping (short: <10 s, medium: 10–20 s, long: >20 s).

Table 3. Effect sizes of error magnitude and attack time grouping (short: <10 s, medium: 10–20 s, long: >20 s) on reliability values.

Metric	Time f^2 [95%CI]	Error magnitude f^2 [95%CI]
DC	0.28 [0.24–0.32]	0.96 [0.95–0.98]
Length	0.30 [0.27–0.34]	0.90 [0.89–0.91]
LPW ratio	0.32 [0.28–0.37]	0.92 [0.90–0.94]
Surface area	0.22 [0.19–0.25]	0.86 [0.84–0.88]
Team spread	0.19 [0.17–0.21]	0.73 [0.71–0.75]
Space control	0.28 [0.25–0.31]	0.71 [0.70–0.73]
$Centroid_x$	0.15 [0.14–0.16]	0.10 [0.09–0.11]
$Centroid_y$	0.17 [0.15–0.19]	0.17 [0.15–0.19]

DC = distance between centroid. LPW ratio = length per width

Team centroid representing the weighted position of players in team was identified as the most reliable metric and demonstrated high reliability values even when positional errors were large and time of analysis was short. The team centroid concept has been applied regularly throughout the literature base and has been identified as relevant in critical situations such as teams reorganising after loss of possession and goal scoring opportunities (Frencken et al. 2012; Sampaio and MaçÃs 2012; Gonçalves et al. 2014; Low et al. 2018; Olthof et al. 2018). In contrast, distance between centroids was identified as the least reliable metric with variance due to positional errors frequently exceeding inherent variance in the metric during the game. Previous research has demonstrated that distance between centroids can cross before teams score in small-sided games (Frencken et al. 2011) and that team centroids are generally coupled with oscillations tightly synchronised (Gonçalves et al. 2014). Siegle and Lames (2013) identified that perturbations in the distance between team centroids were associated with critical moments in the match. However, the findings from the current study highlight that the metric may not be reliable in measurement contexts of large positional errors and when investigating discrete events such as short duration attacks, thereby warranting care in future analyses. Practitioners may instead want to focus on distances between group centroids that have been shown in some instances to have higher variation across game (Gonçalves et al. 2014), and alternatively, synchronisation between team centroids may also provide valuable information (Siegle and Lames 2013), although how the errors in data impact procedures such as the Hilbert transform requires scrutiny.

Space control has recently become a popular metric (Spearman et al. 2017; Fernández et al. 2018) where it is used to evaluate passing ability and decision-making (Filetti et al. 2017; Rein et al. 2017; Spearman et al. 2017). However, similar to the distance between centroids, space control also demonstrated poor reliability in the in silico experiments when calculated with 2 and 4 m errors. This is likely in part due to the dichotomies created through the algorithm. Whereas probabilistic approaches such as the one explored by (Spearman et al. 2017) might provide more reliable measurements, various other approaches have been used to measure dominant regions (Rein et al. 2017) including the integration of player speed in attempts to predict areas on the pitch a player will reach first (Filetti et al. 2017; Brefeld et al. 2019). Further research is required to identify if inclusion of player speed in computations would further decrease reliability in contexts where there are large positional errors.

The present study includes muThe analysis identified strong reliability for the space metrics (surface area and team spread) when calculated with 0.5 and 2 m errors. Surface area and team spread are generally used as measurements of team dispersion and aggregation (Bartlett et al. 2012), with teams commonly expanding when attacking and demonstrating compactness when defending (Duarte et al. 2013; Moura et al. 2016). Clear shifts when transitioning from attacks to defending and the large inherent variance this creates may contribute towards a strong signal, and as a result, the high reliability values were measured. Oscillations in the metrics have also been used by researchers to understand dynamics of transitions, with attempts to measure the coordination of opposing teams surface area and team spread using non-linear analysis methods such as relative phase and vector coding (Moura et al. 2016; Low et al. 2020). Further research is required to investigate the effects of non-linear analyses on reliability and as a result, the accuracy of conclusions that may be drawn.

Across all metrics, the results were consistent in demonstrating that reliability substantively increased when analyses were conducted over longer periods of time. Most metrics demonstrated low reliability values when analyses were made over a period as short as one minute. Whilst research has been conducted analysing spatio-temporal metrics over different fixed periods of time (Duarte et al. 2013; Moura et al. 2016; Clemente et al. 2018; Gonçalves et al. 2018), it is more common to analyse the data across possessions and sequences such as attacks (Bartlett et al. 2012; Duarte et al. 2012; Headrick et al. 2012; Castellano et al. 2013; Barnabé et al. 2016; Laakso et al. 2017). However, the results of the present study show that attacks generally last for short or medium time periods (e.g., <20 s) and as a result, reliability of spatial temporal metrics may be limited even when using measurement technologies generating positional errors as low as 0.5 m. Therefore, researchers and practitioners should remain cautious when analysing spatio-temporal metrics across short possessions.

The present study includes multiple limitations that should be considered when interpreting the findings. The primary limitation of the analysis is the relatively simple error model used for the in silico experiments. The independent Gaussian errors do not fully capture the true properties of positional tracking including relationships between player speed, type of action (e.g., accelerations and change of direction), and position error (Siegle et al. 2013). Accelerations and high-speed actions may be more likely to occur during critical events such as attacks influencing one of the central contexts investigated in this study. As such, the present study should be considered in the context of an initial attempt to provide information regarding reliability of important and increasingly popular metrics, where this information is at present lacking and ultimately required to ensure validity. In addition, more advanced Markov models may be more appropriate to describe positional error and autocorrelation that is likely to occur. As an initial attempt to quantify reliability in this rapidly growing area, the study only considered traditional linear analyses; however, non-linear approaches including relative phase and approximate entropy are frequently applied to spatiotemporal metrics in football. Based on the information reported in this investigation and others, current LPM and video tracking systems appear accurate enough to calculate spatial temporal metrics reliably. Some GPS systems may also be appropriate for use; however, practitioners should remain cautious when using such data. In addition, it was assumed in the present study that very high reliability values were required for subsequent analyses to be valid. However, requirements are likely to be context-dependent and depend on individual judgment. Future research is required to explore the factors that influence reliability incorporating more complex error models with the full range of metrics and analysis approaches used by researchers and practitioners.

Practical applications

The results of the present analysis highlight that researchers and practitioners using spatio-temporal metrics to analyse collective behaviours in football should carefully consider the tracking technologies used to obtain data, the metrics selected and the analysis procedures implemented. A range of metrics including those belonging to distances and numerical relation categories may suffer from unacceptable reliability, particularly when analyses are completed across relatively low samples and analysed across short periods of times. Where researchers and practitioners choose to analyse spatio-temporal metrics across discrete sequences of plays such as attacks, they should consider the suggested bandings used in this study for short, medium, and long sequences as <10 s, 10–20 s and >20 s, respectively. Additionally, individuals should be cautious of using short attacks as these may be unsuitable to draw conclusions based on low reliability. As tracking technology develops, it is likely that more systems will be able to reliably measure a greater range of spatio-temporal metrics over shorter periods of time.

Disclosure statement

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

Funding

The author(s) reported that there is no funding associated with the work featured in this article.

ORCID

Paul Alan Swinton http://orcid.org/0000-0001-9663-0696

References

Aquino R, Guilherme L, Gonçalves C, Henrique L, Vieira P, De L. 2016a. Effects of plyometric training on physiologic, mechanic and neuromuscular variables of long distance runners view project description of the tactical knowledge of training soccer players view project. [Internet]. https://www.researchgate.net/publication/289525744

Aquino RLQT, Cruz Gonçalves LG, Palucci Vieira LH, Oliveira LP, Alves GF, Pereira Santiago PR, Puggina EF. 2016b. Periodization training focused on technical-tactical ability in young soccer players positively affects biochemical markers and game performance. J Strength and Conditioning Res. 30(10):2723–2732. DOI:10.1519/JSC.0000000000001381.

Baptista J, Travassos B, Gonçalves B, Mourão P, Viana JL, Sampaio J. 2020 Jul 1. Exploring the effects of playing formations on tactical behavior and external workload during football small-sided games. J Strength and Conditioning Res. 34(7):2024–2030. DOI:10.1519/JSC.0000000000002445.

Barnabé L, Volossovitch A, Duarte R, Ferreira AP, Davids K. 2016 Aug 1. Age-related effects of practice experience on collective behaviours of football players in small-sided games. Hum Mov Sci. 48: 74–81. DOI:10.1016/j.humov.2016.04.007.

Bartlett R, Button C, Robins M, Dutt-Mazumder A, Kennedy G. 2012. Analysing team coordination patterns from player movement trajectories in soccer: methodological considerations. Int J Perform Anal Sport. 12(2):398–424. DOI:10.1080/24748668.2012.11868607.

Beltrame T, Gois MO, Hoffmann U, Koschate J, Richard X, Hughson L, Linares SN, da Silva Torres R, Catai AM. 2020. Relationship between maximal aerobic power with aerobic fitness as a function of signal-to-noise ratio. Journal of Applied Physiology. 129(3):522–532. DOI:10.6084/m9.figshare.12206654.

Brefeld U, Lasek J, Mair S. 2019 Jan 15. Probabilistic movement models and zones of control. Mach Learn. 108(1):127–147. DOI:10.1007/s10994-018-5725-1.

Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, Skaug H, Mächler M, Bolker B. 2017. GlmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R J. 9(2):378–400. DOI:10.32614/RJ-2017-066.

Buchheit M, Simpson BM. 2017. Player-tracking technology: half-full or half-empty glass? Int J Sports Physiol Perform Human Kinet Publishers Inc. 12:35–41.

Castellano J, Álvarez D, Figueira B, Coutinho D, Sampaio J. 2013. Identifying the effects from the quality of opposition in a Football team positioning strategy. Int J Perform Anal Sport. 13(3):822–832. DOI:10.1080/24748668.2013.11868691.

Castellano J, Silva P, Usabiaga O, Barreira D. 2016 Jun 1. The influence of scoring targets and outer-floaters on attacking and defending team dispersion, shape and creation of space during small-sided soccer games. J Hum Kinet. 50(2):153–163. DOI:10.1515/hukin-2015-0178.

Castellano J, Fernández E, Echeazarra I, Barreira D, Garganta J. 2017. Influencia de la longitud del campo en los comportamientos inter e intra -equipo en jóvenes jugadores de fútbol. Anales de Psicologia. 33(3):486–496. DOI:10.6018/analesps.33.3.271051.

Chung D, Carvalho T, Casanova F, Silva P. 2019 Feb 1. Number of players manipulation effect on space and concentration principles of the game representativeness during football small-sided and conditioned games. J Phys Educ Sport. 19:381–386.

Clemente FM, Couceiro MS, Martins FML, Mendes R. 2013. An online tactical metrics applied to football game. Res J Appl Sci, Eng Technol. 5 (5):1700–1719. DOI:10.19026/rjaset.5.4926.

Clemente FM, Sequeiros JB, Correia A, Serra-Olivares J, González-Víllora S, Silva F, Martins FM. 2018 Jan 1. How dots behave in two different pitch sizes? Analysis of tactical behavior based on position data in two soccer field sizes. RICYDE: Revista Int de Ciencias del Deporte. 14(51):16–28.

Cummins C, Orr R, O'Connor H, and West C. 2013. Global positioning systems (GPS) and microtechnology sensors in team sports: a systematic review. Vol. 43, Sports Medicine. p. 1025–1042 doi:10.1007/s40279-013-0069-2.

Duarte R, Araújo D, Davids K, Travassos B, Gazimba V, Sampaio J. 2012 May. Interpersonal coordination tendencies shape 1-vs-1 sub-phase performance outcomes in youth soccer. J Sports Sci. 30(9):871–877. DOI:10.1080/02640414.2012.675081.

Duarte R, Araújo D, Folgado H, Esteves P, Marques P, Davids K. 2013 Feb 1. Capturing complex, non-linear team behaviours during competitive football performance. J Syst Sci Complexity. 26(1):62–72. DOI:10.1007/s11424-013-2290-3.

Fernández J, Barcelona FC, Fernandez J, Bornn L Wide Open Spaces: a statistical technique for measuring space creation in professional soccer [Internet]. 2018. http://www.lukebornn.com/sloan/space_occupation_1.mp4

Filetti C, Ruscello B, D'Ottavio S, Fanelli V. 2017 Jun 1. A study of relationships among technical, tactical, physical parameters and final outcomes in elite soccer matches as analyzed by a semiautomatic video tracking system. Percept Mot Skills. 124(3):601–620. DOI:10.1177/0031512517692904.

Frencken WGP, Lemmink KAPM, Delleman NJ. 2010 Nov. Soccer-specific accuracy and validity of the local position measurement (LPM) system. J Sci Med Sport. 13(6):641–645. DOI:10.1016/j.jsams.2010.04.003.

Frencken W, Lemmink K, Delleman N, Visscher C. 2011 Jul. Oscillations of centroid position and surface area of soccer teams in small-sided games. Eur J Sport Sci. 11(4):215–223. DOI:10.1080/17461391.2010.499967.

Frencken W, de Poel H, Visscher C, Lemmink K. 2012 Aug. Variability of inter-team distances associated with match events in elite-standard soccer. J Sports Sci. 30(12):1207–1213. DOI:10.1080/02640414.2012.703783.

Frencken W, van der Plaats J, Visscher C, Lemmink K. 2013 Feb 1. Size matters: pitch dimensions constrain interactive team behaviour in soccer. J Syst Sci Complexity. 26(1):85–93. DOI:10.1007/s11424-013-2284-1.

Frias T. 2014. Man-to-man or zone defense measuring team dispersion behaviors in small-sided soccer games. Trends Sport Sci. 21(3):135

Gonçalves BV, Figueira BE, Maçãs V, Sampaio J. 2014 Jan. Effect of player position on movement behaviour, physical and physiological performances during an 11-a-side football game. J Sports Sci. 32(2):191–199. DOI:10.1080/02640414.2013.816761.

Gonçalves B, Marcelino R, Torres-Ronda L, Torrents C, Sampaio J. 2016 Jul 17. Effects of emphasising opposition and cooperation on collective movement behaviour during football small-sided games. J Sports Sci. 34(14):1346–1354. DOI:10.1080/02640414.2016.1143111.

Gonçalves B, Coutinho D, Travassos B, Folgado H, Caixinha P, Sampaio J. 2018 Jul 1. 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.

Gonçalves B, Coutinho D, Exel J, Travassos B, Lago C, Sampaio J, Boullosa D. 2019 Aug 1. Extracting spatial-temporal features that describe a team match demands when considering the effects of the quality of opposition in elite football. PLoS ONE. 14(8):e0221368. DOI:10.1371/journal.pone.0221368.

Hämmerer D, Li SC, Völkle M, Müller V, Lindenberger U. 2013. A lifespan comparison of the reliability, test-retest stability, and signal-to-noise ratio of event-related potentials assessed during performance monitoring. Psychophysiology. 50(1):111–123. DOI:10.1111/j.1469-8986.2012.01476.x.

Headrick J, Davids K, Renshaw I, Araújo D, Passos P, Fernandes O. 2012. Proximity-to-goal as a constraint on patterns of behaviour in attacker–defender dyads in team games. J Sports Sci. 30(3):247–253. DOI:10.1080/02640414.2011.640706.

Laakso T, Travassos B, Liukkonen J, Davids K. 2017 Aug 1. Field location and player roles as constraints on emergent 1-vs-1 interpersonal patterns of play in football. Hum Mov Sci. 54: 347–353. DOI:10.1016/j.humov.2017.06.008.

Laakso T, Davids K, Liukkonen J, Travassos B. 2019. Interpersonal dynamics in 2-vs-1 contexts of football: the effects of field location and player roles. Front Psychol. 10(JUL). DOI:10.3389/fpsyg.2019.01407.

Lacome M, Peeters A, Mathieu B, Marrier B, Carling C, Piscione J. 2019. Can we use GPS for assessing sprinting performance in rugby sevens? A concurrent validity and between-device reliability study. Biol Sport. 36(1):25–29. DOI:10.5114/biolsport.2018.78903.

Leser R, Moser B, Hoch T, Stögerer J, Kellermayr G, Reinsch S, Baca A. 2015. Expert-oriented modelling of a 1vs1-situation in football. Int J Perform Anal Sport. 15(3):949–966. DOI:10.1080/24748668.2015.11868843.

Linke D, Link D, Lames M. 2018 Jul 1. Validation of electronic performance and tracking systems EPTS under field conditions. PLoS ONE. 13(7): e0199519. DOI:10.1371/journal.pone.0199519.

Linke D, Link D, Lames M. 2020. Football-specific validity of TRACAB's optical video tracking systems. PLoS One. 15(3):e0230179. DOI:10.1371/journal.pone.0230179.

Low B, Boas GV, Meyer L, Lizaso E, Hoitz F, Leite N, Low B, Boas GV, Meyer L, Lizaso E, Hoitz F, Leite N, Low B, Boas GV, Meyer L, Lizaso E, Hoitz F, Leite N, and Gonçalves B. 2018. Exploring the effects of deep-defending vs high-press on footballers' tactical behaviour, physical and physiological performance: a pilot study. Motriz Revista de Educacao Fisica. 24(2):e1018171. doi:10.1590/S1980-6574201800020009

Low B, Coutinho D, Gonçalves B, Rein R, Memmert D, and Sampaio J. 2020. A systematic review of collective tactical behaviours in football using positional data. Vol. 50, Sports Medicine. Springer. p. 343–385 doi:10.1007/s40279-019-01194-7.

Menuchi MRTP, Moro ARP, Ambrósio PE, Pariente CAB, and Araújo D. 2018. Spatiotemporal Constraints and Age on the Interactions of Soccer Players when Competing for Ball Possession [Internet]. ©J Sports Sci Med. 17(3): 379–391. http://www.jssm.org

Moura FA, Martins LEB, de Oliveira Anido R, de Barrosr RML, Cunha SA, De Barros RML. 2012 Mar. Quantitative analysis of Brazilian football players' organisation on the pitch. Sports Biomech. 11(1):85–96. DOI:10.1080/14763141.2011.637123.

Moura FA, Martins LEB, Anido RO, Ruffino PRC, Barros RML, Cunha SA. 2013 Oct. A spectral analysis of team dynamics and tactics in Brazilian football. J Sports Sci. 31(14):1568–1577. DOI:10.1080/02640414.2013.789920.

Moura FA, van Emmerik REA, Santana JE, Martins LEB, de Barros RML, Cunha SA, van Emmerik REA, Barros RMLD. 2016. Coordination analysis of players' distribution in football using cross-correlation and vector coding techniques. J Sports Sci. 34(24):2224–2232. DOI:10.1080/02640414.2016.1173222.

Nakagawa S, Schielzeth H. 2012. The mean strikes back: mean-variance relationships and heteroscedasticity. Trends Ecol Evol. 27(9):2. DOI:10.1016/j.tree.2012.04.003.

Ogris G, Leser R, Horsak B, Kornfeind P, Heller M, Baca A. 2012 Oct. Accuracy of the LPM tracking system considering dynamic position changes. J Sports Sci. 30(14):1503–1511. DOI:10.1080/02640414.2012.712712.

Olthof SBH, Frencken WGP, Lemmink KAPM. 2015 Jun 1. The older, the wider: on-field tactical behavior of elite-standard youth soccer players in small-sided games. Hum Mov Sci. 41: 92–102. DOI:10.1016/j.humov.2015.02.004.

Olthof SBH, Frencken WGP, Lemmink KAPM. 2018 Jul 18. Match-derived relative pitch area changes the physical and team tactical performance of elite soccer players in small-sided soccer games. J Sports Sci. 36(14):1557–1563. DOI:10.1080/02640414.2017.1403412.

Olthof SBH, Frencken WGP, Lemmink KAPM. 2019. When something is at stake: differences in soccer performance in 11 vs. 11 during official matches and training games. J Strength and Conditioning Res. 33(1):167–173. [Internet]. www.nsca.com

Palucci Vieira LH, Aquino R, Moura FA, Barros RMLD, Arpini VM, Oliveira LP, Bedo BLS, Santiago PRP. 2019 Aug 1. Team dynamics, running, and skill-related performances of Brazilian U11 to professional soccer players during official matches. J Strength and Conditioning Res. 33(8):2202–2216. DOI:10.1519/JSC.0000000000002577.

Praça GM, Folgado H, de Andrade AGP, Greco PJ. 2016. Comportamento tático coletivo em Pequenos Jogos no Futebol: influência de jogadores adicionais. Braz J Kinanthropometry and Human Perform. 18(1):62–71. DOI:10.5007/1980-0037.2016v18n1p62.

Rein R, Raabe D, Memmert D. 2017 Oct 1. "Which pass is better?" Novel approaches to assess passing effectiveness in elite soccer. Hum Mov Sci. 55: 172–181. DOI:10.1016/j.humov.2017.07.010.

Ryan S, Pacecca E, Tebble J, Hocking J, Kempton T, Coutts AJ. 2020. Measurement characteristics of athlete monitoring tools in professional Australian football. Int J Sports Physiol Perform. 15(4):457–463. DOI:10.1123/ijspp.2019-0060.

Sampaio J, MaçÃs V. 2012. Measuring tactical behaviour in football. Int J Sports Med. 33(5):395–401. DOI:10.1055/s-0031-1301320.

Santos PM, Lago-Penas C. 2019. Defensive positioning on the pitch in relation with situational variables of a professional football team during regaining possession. Human Mov. 20(2):50–56. DOI:10.5114/hm.2019.81019.

Sarmento H, Clemente FM, Araújo D, Davids K, McRobert A, Figueiredo A. 2018 Apr 1. What performance analysts need to know about research trends in association football (2012–2016): a systematic review. Sports Med. 48(4):799–836. DOI:10.1007/s40279-017-0836-6.

Shafizadeh M, Davids K, Correia V, Wheat J, Hizan H. 2016. Informational constraints on interceptive actions of elite football goalkeepers in 1v1 dyads during competitive performance [Internet]. J Sports Sci. 34 (17):1596–1601. http://shura.shu.ac.uk/information.html

Siegle M, Stevens T, Lames M. 2013 Jan. Design of an accuracy study for position detection in football. J Sports Sci. 31(2):166–172. DOI:10.1080/02640414.2012.723131.

Siegle M, Lames M. 2013 Feb 1. Modeling soccer by means of relative phase. J Syst Sci Complexity. 26(1):14–20. DOI:10.1007/s11424-013-2283-2.

Silva P, Duarte R, Sampaio J, Aguiar P, Davids K, Araújo D, Garganta J. 2014a Dec 14. Field dimension and skill level constrain team tactical behaviours in small-sided and conditioned games in football. J Sports Sci. 32 (20):1888–1896. DOI:10.1080/02640414.2014.961950.

Silva P, Travassos B, Vilar L, Aguiar P, Davids K, Araújo D, Garganta J. 2014b Sep 5. Numerical relations and skill level constrain co-adaptive behaviors of agents in sports teams. PLoS One. 9(9):e107112. DOI:10.1371/journal.pone.0107112.

Spearman W, Basye AT, Basye A, Dick G, Hotovy R, Pop P Physics-based modeling of pass probabilities in soccer [Internet]. 2017. https://www.researchgate.net/publication/315166647

Steiner S, Rauh S, Rumo M, Sonderegger K, Seiler R. 2019 Jun 1. Outplaying opponents—a differential perspective on passes using position data. Ger J Exercise and Sport Res. 49(2):140–149. DOI:10.1007/s12662-019-00579-0.

Travassos B, Vilar L, Araújo D, McGarry T. 2014. Tactical performance changes with equal vs unequal numbers of players in small-sided football games. Int J Perform Anal Sport. 14(2):594–605. DOI:10.1080/24748668.2014.11868745.