Exploring the determinants of success in different clusters of ball possession sequences in soccer

Murilo Merlin[a*], Sergio Augusto Cunha[a], Felipe Arruda Moura[b], Ricardo da Silva Torres[c], Bruno Gonçalves[d] and Jaime Sampaio[d]

[a]School of Physical Education, University of Campinas, Campinas, Brazil; [b]Laboratory of Applied Biomechanics, State University of Londrina, Londrina, Brazil; [c]Department of ICT and Natural Sciences, Faculty of Information Technology and Electrical Engineering, NTNU – Norwegian University of Science and Technology, Ålesund, Norway; [d]Research Center in Sports Sciences, Health Sciences and Human Development (CIDESD), CreativeLab Research Community, University of Tras-os-Montes and Alto Douro, Vila Real, Portugal; e Portugal Football School, Portuguese Football Federation, Oeiras, Portugal

*CONTACT: Murilo Merlin / E-mail: murilomerlin7@gmail.com / Faculdade de Educação Física, Universidade Estadual de Campinas, Av. Érico Veríssimo, 701, Campinas, São Paulo, CEP 13083-851, Brasil.

Exploring the determinants of success in different clusters of ball possession sequences in soccer

ABSTRACT

The purpose of this study was two-step: classify ball possession (BP) according to the duration and number of passes; identify which tactical variables most discriminate the different BP. We obtained 527 BPs from four official matches of the Brazilian Soccer Championship 2016. Forty-one 'notational', 'space occupation', and 'displacement synchronization' predictor variables were used. The BPs were classified into three groups: short (11.07 ± 4.49s, 1.93 ± 0.99 passes), medium (26.83 ± 7.33s, 5.41 ± 1.84 passes), long (55.50 ± 14.97s, 12.11 ± 4.61 passes). Discriminant analysis identified the five most relevant variables to describe each group: coefficient of variation (CV) of the defensive team's synchronization-Y, CV defensive team´s synchronization-X, successful pass last third, CV distance between offensive team's centroid and target, mean of the offensive team's width. The approach highlights important variables and could benefit the description of offensive and defensive game sequences to provide precise knowledge on the process.

Keywords: ball possession; tactical; multivariate; soccer

Introduction

Ball possession (BP) is the consequence of interactions determined by contextual factors, such as quality of opponent, tactical configuration, match status, or venue of the match (Link, Hoernig, Nassis, Laughlin, & Witt, 2017). Tactically, controlling the ball possession as much possible consists of a substantial set of on-ball and off-ball actions to generate scoring chances. Some of these actions are associated with game principles like creating numerical superiority or promoting disorder on the opponents' defense, but most importantly, generating and occupying spaces (Fernandez & Bornn, 2018).

Although BP is a complex phenomenon whose success depends on the combination of many variables, most research insist in an attempt to establish a cause-effect relationship, ie, how BP's time influences performance indicators such as shots and goals or performance across the season (Collet, 2013; Hughes & Franks, 2005). Besides that, literature studies have explored others properties of BP, considering aspects such as passing frequency, pitch zones where the ball moves, passing characteristics and match status (Cintia, Giannotti et al., 2015; Lago & Martín, 2007; P. D. Jones, 2004; Paixão et al., 2015a).

In our viewpoint, more important that to relate BP's properties to performance indicators, is identify and describe collective behaviors that help to maintain BP and perform passes, considering their relevance to the match.

For this topic, recent research has proposed several variables that compose collective movement behaviour (Memmert, Lemmink, & Sampaio, 2016). When the analysis is focused on the dynamics of space occupation, variables such as the coverage area or effective playing space (Moura et al., 2012), length, width, and measures around the centroid (Folgado, Lemmink, Frencken, & Sampaio, 2014; Coutinho et al., 2019) are widely used. Besides that, several non-linear processing techniques have been used to improve the performance analysis process. For example, approximate entropy (ApEn) appears to provide information about the regularity of certain behaviour in soccer games and seems to be associated with adaptation during training interventions (Sampaio & Maçãs, 2012), critical moments of the game (Aguiar, Gonçalves, Botelho, Duarte, & Sampaio, 2017), or interpersonal game distances (Gonçalves et al., 2016). Complementarily to this structure of variability, the coefficient of variation (CV) has also been used to measure the magnitude of the variability of a given behaviour across time (Gonçalves et al., 2017; Lorenzo-Martínez et al., 2019; Castillo, et al., 2019).

Non-linear processing techniques have also been used to identify coordination patterns in tactical behaviour analyses. Several studies have shown that movement synchronization is linked to tactical performance (Folgado, Duarte, Fernandes, & Sampaio, 2014; Folgado, Gonçalves, Sampaio, Folgado, & Gonçalves, 2017), with consequences on the external and internal workload demands (Folgado, Duarte, Marques, Gonçalves, & Sampaio, 2018).

Considering the previous arguments, a multivariate approach based on metrics that describe collective behaviors in BP sequences could provide a more holistic model of this phenomenon in soccer matches. Within this topic, the outcomes would benefit from descriptions of the offensive and defensive game sequences to provide precise knowledge on the process. In addition, there are few studies on ball possession that describe collective tactical behaviours that determine the team ability to maintain ball possession. Thus, the purpose of this study was two-step: i) classify ball possession sequences according to the duration and number of passes; ii) identify which tactical variables most discriminate the different ball possession sequences, as classified in the previous step.

Methods

Data collection and sample

The Ethics Committee of the Campinas State University approved this research. The sample of this study corresponds to 527 ball possession (BP) sequences obtained from four first division official matches of the Brazilian Soccer Championship 2016.

The matches were recorded by two digital cameras (HDR-CX405, Sony), HD resolution, acquisition frequency of 15Hz, commonly used in collective tactical analysis (Rico-gonzález, Arcos, Nakamura, Arruda, & Pino-ortega, 2019). Subsequently, a semiautomatic tracking system was used to obtain the players' 2D positional data using the software DVideo (Pascual, Leite, & Barros, 2002; Figueroa, Leite, & Barros, 2006). The 2D coordinates of each player were defined as Xp(t) and Yp(t), where t represents each instant of time. The X and Y axes represent length and width of the pitch respectively. A Butterworth third-order low-pass digital filter with a cut-off frequency of 0.4 Hz was used as an external filter according to previous study recommendations (Barros et al., 2007). DVideo software has an automatic tracking rate of 94% of the processed frames, an average error of 0.3 m for the determination of player position, and an average error of 1.4% for the distance covered (Figueroa, Leite, & Barros, 2006). Notational analysis was performed by an experienced operator to register the technical actions of each player, synchronized with the positioning data.

Ball possession sequences

Each ball possession started when any player controlled the ball through the successful execution of a technical action, such as a pass, interception or tackle, and restarting play, such as a free kick, throw-in, corner kick, and goal kick. When the game stopped for less than 15 seconds and the ball remained with the same team, it was considered the same BP sequence. This decision was made since the match dynamics of player positioning were not influenced. BP sequences of less than four seconds were excluded (to fulfil the nonlinear computation requirements). BP that did not contain at least one successful pass were also excluded.

Variables

Forty-one variables were computed and classified into three groups: notational, space occupation, and displacement synchronization (Table 1). Dynamic variables were analysed using the absolute values (mean), normalized approximate entropy (ApEn), and coefficient of variation (CV). ApEn is a nonlinear measure that quantifies the regularity in complex system behaviors (Pincus, 1991). For this study, we decided to compute the normalized entropy, a non-modified measure of regularity derived from the original ApEn, which is less dependent on time series length (Fonseca, Milho, Passos, Araújo, & Davids, 2012). Coefficient of variation (CV) values ((standard deviation/mean)×100) were used to verify the magnitude of variability of the time series.

The displacement synchronization variables consisted of the percentage of time that inter-player displacements were synchronized, calculated using the vector coding technique (Sparrow, Donovan, Van Emmerik, & Barry, 1987) and recently applied to investigate player behaviour during tennis matches (Pereira, van Emmerik, Misuta, Barros, & Moura, 2017). The technique consists of calculating the angle (θ) formed by the relative motion between two oscillators in two consecutive coordinates of a given time series. The coupling angle represents an instantaneous spatial relationship between two players (dyad) in relation to the axes (X and Y). The coupling was considered as in-phase when the angle was at 45 ° or 225 ° (positive diagonal). Thus, the intervals $22.5° \leq \theta < 67.5°$ and $202.5° \leq \theta < 247.5°$ were chosen to assume an in-phase synchronization between two players. The synchronization percentage for each dyad was calculated for each team (in possession and without possession), in each ball possession sequence. The mean values of the percentage (% mean) of all the dyads were used to represent the mean of team synchronization and the CV (based on the % mean of all dyads) was calculated to indicate the variability between the dyads, i.e., if there was homogeneous behaviour of the team. All these procedures were performed for the X (longitudinal) and Y (lateral) axes of the pitch reference. Space occupation and synchronization variables are shown in Figures 1a and 1b, respectively. Data processing was performed in Matlab®2017(Mathworks Inc., Natick, MA, USA).

Statistical analysis

A two-step cluster with log-likelihood as the distances measure and the Schwartz's Bayesian criterion was used to classify the ball possession sequences into the different groups, according to the time of possession and number of successful passes. Afterwards, a stepwise fisher´s discriminant analysis (FDA) was conducted to identify which variables best discriminate the previously obtained clusters. At each step, the variable that minimized the overall Wilks' Lambda was entered in the model. A minimum partial F (Fisher) value (3.84) to enter and maximum partial F value (2.71) to remove was used. Validation of discriminant models was conducted using the leaveone-out method of cross-validation. Was applied One-way ANOVA was used to compare the twelve selected variables into different groups (short, medium, and long ball possession sequences). Subsequently, the Bonferroni post-hoc test was utilized to identify pairwise differences. Statistical significance was set at 0.05 and the statistical analysis was carried out in IBM SPSS Statistics for Windows (Armonk, NY: IBM Corp). Complementarily, was observed the standardized mean differences and respective 95% confidence limits (CL), were also computed as magnitude of observed differences, effect size (Cohen´s *d*) and thresholds were: <0.2, trivial; 0.6, small; 1.20, moderate; 2.0, large; and >2.0, very large (Hopkins et al., 2009).

Results

The 527 ball possession sequences (BP) were classified into three different groups according to the time duration and number of successful passes: cluster 1 (short possessions n=295 or 55.8%, 11.07 ± 4.49s, 1.93 ± 0.99 successful passes), cluster 2 (medium possessions n=179 or 34%, 26.83 ± 7.33s, 5.41 ± 1.84 successful passes), and cluster 3 (long possessions n=53 or10.3%, 55.50 ± 14.97s, 12.11 ± 4.61 successful passes).

The stepwise fisher´s discriminant analysis (FDA) identified the most relevant variables to describe each cluster. The model consisted of two discriminant functions, with function 1 representing 95.8% of the total variance and function 2 representing 4.2%. The canonical correlations of functions 1 and 2 were, respectively, 0.83 and 0.30, with both functions being statistically significant (p <0.0001), (Wilks' Lambda = 0.27 and 0.91 for functions 1 and 2, respectively). The model presented a total of 81.6% of the original grouped cases classified correctly. Table 2 presents the descriptive analysis for each variable, for the three clusters, as well as the structure coefficients (SC) for each function.

The variables that contributed most to the classification of the BP into function 1, in order of importance were: CV of the defensive synchronization-Y (SC = 0.58), CV of the defensive synchronization-X (SC = 0.42), successful pass last third, CV of the distance between offensive centroid and target (SC = 0.34), and mean of the offensive width (SC = 0.33). The remaining seven variables were: centroid progression, % mean of the offensive synchronization-X, CV of the offensive synchronization-X, % mean of the defensive synchronization-X, mean of the defensive length, and mean of the distance between offensive centroid and target.

Figure 2 represents the canonical discriminant function by distribution of the possession linked to cluster centroids, based on the discriminant scores represented by the X axis (function 1) and the Y axis (function 2).

Discussion

The purpose of this study was two-step: i) classify ball possession sequences according to the duration and number of passes; ii) identify which tactical variables most discriminate the different ball possession sequences, as classified in the previous step. In the first step, the cluster analysis classified the ball possession (BP) into three groups, short, medium and long duration. This classification allowed identify, describe and compare the collective tactical behavior to both teams, in offensive and defensive phase. For this, in the second step we use FDA to highlight, between forty-one tactical variables, the most relevant that better describe these three clusters. Five variables were highlighted: coefficient of variation (CV) of the defensive team’s synchronization-Y, CV defensive team´s synchronization-X, successful pass last third, CV distance between offensive team’s centroid and target, mean of the offensive team’s width. The findings provided accurate tactical characterization to offensive and defensive team´s in the short, medium and long BP sequences and therefore suggest collective behaviors that help to maintain BP and perform passes, which is one of the challenges of the offensive phase of the matches.

In relation to the ball possession clusters identified, Aguiar et al. (2017) also classified BP using cluster analysis, however found two distinct groups, short and long, and the criterion for separation was based on centroid approximate entropy measurements. Jones et al. (2004) proposed three categories of ball possession durations, 3-10s, 10-20s, and more than 20s to investigate the relation with match status. Other studies with BP did not review the time duration or the number of passes and usually compared short and long sequences (Collet, 2013; da Mota et al., 2015; Yiannakos & Armatas, 2017).

In the present study, the short ball possession duration was characterized by lower successful passes in the last third, high CV of defensive team's synchronization in relation to X-axis and Y-axis, lower CV of distance between offensive team's centroid and target, and lower mean offensive team width. On the other hand, when we analysed the long ball possession duration, we observed more successful passes in the last third of the pitch, smaller CV of defensive team's synchronization in relation to X-axis and Y-axis, higher CV of distance between offensive team's centroid and target, and higher mean of the offensive team width. The medium ball possession duration presented intermediate values for the five variables.

The successful passes in the last third was the only notational variable highlighted. Displacement synchronization variables demonstrated importance for classification of the cluster, especially through the CV values of the defending team. These variables represent the variability of the percentage values of all team dyads. That is, the higher the CV, the more heterogenic the behaviour of the dyadic relations during the time series, as observed in short ball possessions. Otherwise, when dyads present similar behaviours between them, the CV values decrease, characterized in longer ball possessions. It is probable this behaviour is associated with the transition phases and stabilization in the possessions, i.e., when there is loss of the ball, the defensive team reorganizes strategically into its new tactical functions, changing the dynamics of space occupation during this transition. In short ball possessions, characterized as a mean of 11.7 s duration, there is no stabilization moment, or the transition phase is predominant, reflecting in the high CV of synchronization in relation to the X and Y axes. In the long possessions, there is also a transition phase, following a long period of stabilization, which probably explains the lower CV. These behaviours were conceptually identified by Hewitt et al. (2016), who generally describe the game as moments of frenetic attack to create imbalances in the opponent and moments of homeostasis, with rapid reorganization towards control and stability between the teams. Moura et al. (2013) also describe similar behaviour, but through the dynamics of the team occupying area, assigning higher values, based on spectral analysis, at the moment of the game where teams change ball possession rapidly, i.e., short possessions.

The other two highlighted variables belong to the 'space occupation' group. The CV of the distance between the offensive team's centroid and target indicated greater variability in longer ball possessions. It is probable the greater mobility of the team in possession exploring the pitch favoured the passes performed and control of the ball, as well as the width of the offensive team, which was higher in long ball possessions. It seems clear that teams adopting wider pitch space occupation and mobility favoured BP. Mobility and width are two of the five most important offensive principles proposed by Ouellette (2004). According to Clemente et al. (2013), the movements of players should extend to use the effective playing space by increasing the dispersion of players during the offensive phase. This behaviour makes it easier to attract defensive players to non-vital zones (e.g., lateral zones), thereby removing them from the vital zones (i.e., the middle zones). Clearly, it is essential to analyse offensive and defensive behaviour from the interaction between teams, not just from a single perspective, as proposed by Fernandez-Navarro et al. (2016).

In summary, ball possession sequences were classified into three clusters based on the time possession and number of successful passes: short, medium and long duration. The discriminant analysis highlighted five most important variables to describe each cluster, and thus, these should be observed with more attention by coaches and sports scientists. Long ball possessions durations were characterized by more homogeneous behavior of the defending team in relation to displacements in lateral and longitudinal directions. There are few studies related to this phenomenon and therefore, their association with the micro-level relations among teammates should be further explored. Completely, higher width and mobility of the offensive team in long ball possession reinforcing some principles of offensive game advocated by experts, with the advantage of having been quantified and not only subjectively identified. This study used a limited sample based on Brazilian Soccer Championship and therefore should not be conclusive. The approach based on a multivariate model, using metrics recently proposed by research in performance analysis, allowed holistic analysis of the phenomena and provided accurate knowledge.

References

Aguiar, M., Gonçalves, B., Botelho, G., Duarte, R., & Sampaio, J. (2017). Regularity of interpersonal positioning discriminates short and long sequences of play in small-sided soccer games. *Science and Medicine in Football*, *00*(00), 1–7. https://doi.org/10.1080/24733938.2017.1353220

Barros, R., Misuta, M., Menezes, R., Figueroa, P., Moura, F., Cunha, S., … Leite, N. (2007). Analysis of the distances covered by first division Brazilian soccer players obtained with an automatic tracking method. *Journal of Sports Science and Medicine*, *6*, 233–242.

Castillo, D., Raya-gonzález, J., Clemente, F. M., Yanci, J., Castillo, D., … Clemente, F. M. (2019). The influence of youth soccer players ' sprint performance on the different sided games ' external load using GPS devices. *Research in Sports Medicine*. https://doi.org/10.1080/15438627.2019.1643726

Cintia, P., Giannotti, F., Pappalardo, L., Pedreschi, D., & Malvaldi, M. (2015). The harsh rule of the goals: Data-driven performance indicators for football teams. *Proceedings of the 2015 IEEE International Conference on Data Science and Advanced Analytics, DSAA 2015*. https://doi.org/10.1109/DSAA.2015.7344823

Clemente, F. M., Cauceiro, M. S., Martins, F. M., Ivanova, M. O., & Mendes, R. (2013). Activity Profiles Soccer During the 2010 World Cup. *Journal of Human Kinetics.*, *38*, 201–211.

Collet, C. (2013). The possession game? A comparative analysis of ball retention and team success in European and international football, 2007-2010. *Journal of Sports Sciences*, *31*(2), 123–136. https://doi.org/10.1080/02640414.2012.727455

Coutinho, D., Gonçalves, B., Santos, S., Travassos, B., Del, P., Sampaio, J., Travassos, B. (2019). Effects of the pitch configuration design on players ' physical performance and movement behaviour during soccer small-sided games. *Research in Sports Medicine*, *27*(3), 298–313. https://doi.org/10.1080/15438627.2018.1544133

da Mota, G. R., Thiengo, C. R., Gimenes, S. V., & Bradley, P. S. (2015). The effects of ball possession status on physical and technical indicators during the 2014 FIFA World Cup Finals. *Journal of Sports Sciences*, *0414*(January 2016), 1–8. https://doi.org/10.1080/02640414.2015.1114660

Fernandez-navarro, J., Fradua, L., Zubillaga, A., Ford, P. R., Mcrobert, A. P., Fradua, L., Fernandez-navarro, J. (2016). Attacking and defensive styles of play in soccer : analysis of Spanish and English elite teams elite teams, *0414*(April). https://doi.org/10.1080/02640414.2016.1169309

Fernandez, J., & Bornn, L. (2018). Wide Open Spaces : A statistical technique for measuring space creation in professional soccer. *MIT Sloan Sports Analytics Conferece*, (March), 1–19.

Figueroa, P. J., Leite, N. J., & Barros, R. M. L. (2006). Background recovering in outdoor image sequences : An example of soccer players Background recovering in outdoor image sequences : An example of soccer players segmentation, (November 2015). https://doi.org/10.1016/j.imavis.2005.12.012

Folgado, H., Duarte, R., Fernandes, O., & Sampaio, J. (2014). Competing with lower level opponents decreases intra-team movement synchronization and time-motion demands during pre-season soccer matches. *PLoS ONE*, *9*(5). https://doi.org/10.1371/journal.pone.0097145

Folgado, H., Duarte, R., Marques, P., Gonçalves, B., & Sampaio, J. (2018). Exploring how movement synchronization is related to match outcome in elite professional football. *Science and Medicine in Football*, *2*(2), 101–107.

https://doi.org/10.1080/24733938.2018.1431399

Folgado, H., Gonçalves, B., Sampaio, J., Folgado, H., & Gonçalves, B. (2017). Positional synchronization affects physical and physiological responses to preseason in professional football (soccer) responses to preseason in professional football (soccer). *Research in Sports Medicine*, *26*(1), 51–63. https://doi.org/10.1080/15438627.2017.1393754

Folgado, H., Lemmink, K. A. P. M., Frencken, W., & Sampaio, J. (2014). Length, width and centroid distance as measures of teams tactical performance in youth football. *European Journal of Sport Science*, *14*(sup1), S487–S492. https://doi.org/10.1080/17461391.2012.730060

Fonseca, S., Milho, J., Passos, P., Araújo, D., & Davids, K. (2012). Approximate entropy normalized measures for analyzing social neurobiological systems. *Journal of Motor Behavior*, *44*(3), 179–183. https://doi.org/10.1080/00222895.2012.668233

Gonçalves, B., Coutinho, D., Santos, S., Lago-Penas, C., Jiménez, S., & Sampaio, J. (2017). Exploring team passing networks and player movement dynamics in youth association football. *PLoS ONE*, *12*(1), 1–13. https://doi.org/10.1371/journal.pone.0171156

Gonçalves, B., Marcelino, R., Torres-Ronda, L., Torrents, C., & Sampaio, J. (2016). Effects of emphasising opposition and cooperation on collective movement behaviour during football small-sided games. *Journal of Sports Sciences*, *34*(14), 1346–1354. https://doi.org/10.1080/02640414.2016.1143111

Hewitt, A., Greenham, G., & Norton, K. (2016). Game style in soccer: what is it and can we quantify it? *International Journal of Performance Analysis in Sport*, (April), 355–372.

Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive Statistics for Studies in Sports Medicine and Exercise Science, (21), 3–12. https://doi.org/10.1249/MSS.0b013e31818cb278

Hughes, M., & Franks, I. (2005). Analysis of passing sequences, shots and goals in soccer. *Journal of sports sciences*, *23*(5), 509–514. https://doi.org/10.1080/02640410410001716779

Lago, C., & Martín, R. (2007). Determinants of possession of the ball in soccer. *Journal of sports sciences*, *25*(9), 969–974. https://doi.org/10.1080/02640410600944626

Link, D., Hoernig, M., Nassis, G., Laughlin, M., & Witt, J. de. (2017). Individual ball possession in soccer. *Plos One*, *12*(7), e0179953. https://doi.org/10.1371/journal.pone.0179953

Lorenzo-martínez, M., Rey, E., Padrón-cabo, A., Rey, E., The, A. P., Rey, E. (2019). The effect of age on between-match physical performance variability in professional soccer players variability in professional soccer players. *Research in Sports Medicine*. https://doi.org/10.1080/15438627.2019.1680985

Memmert, D., Lemmink, K. A. P. M., & Sampaio, J. (2016). Current Approaches to Tactical Performance Analyses in Soccer Using Position Data. *Sports Medicine*, 1–10. https://doi.org/10.1007/s40279-016-0562-5

Moura, F., Martins, L. E., Anido, R., Barros, R., & Cunha, S. (2012). Quantitative analysis of Brazilian football players' organisation on the pitch. *Sports Biomechanics*, *11*(1), 85–96.

Moura, F., Martins, L. E., Anido, R., Ruffino, P. R., Barros, R., & Cunha, S. (2013). A spectral analysis of team dynamics and tactics in Brazilian football. *Journal of Sports Sciences*, 37–41.

Ouellette, J. (2004). Principles of Play for Soccer. *Strategies*, *17*(October 2014), 3.

https://doi.org/10.1080/08924562.2004.10591082

P. D. Jones, N. J. and S. D. M. D. (2004). Possession as a performance indicator in soccer. *Department of Sports Science*, (August), 98–102. https://doi.org/10.1017/CBO9781107415324.004

Paixão, P., Sampaio, J., Almeida, C. H., & Duarte, R. (2015). How does match status affects the passing sequences of top-level European soccer teams ? *International Journal of Performance Analysis in Sport*, *15*(1), 229–240.

Pascual, F., Leite, N., & Barros, R. (2002). A flexible software for tracking of markers used in human motion analysis. *Computer Methods and Programs in Biomedicine.*, *72*, 155–165.

Pereira, T. J. C., van Emmerik, R. E. A., Misuta, M. S., Barros, R. M. L., & Moura, F. A. (2017). Interpersonal coordination analysis of tennis players from different levels during official matches. *Journal of Biomechanics*, *67*, 106–113. https://doi.org/10.1016/j.jbiomech.2017.11.036

Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences*, *88*(6), 2297–2301. https://doi.org/10.1073/pnas.88.6.2297

Rico-gonzález, M., Arcos, A. L., Nakamura, F. Y., Arruda, F., & Pino-ortega, J. (2019). The use of technology and sampling frequency to measure variables of tactical positioning in team sports : a systematic review. *Research in Sports Medicine*, *00*(00), 1–14. https://doi.org/10.1080/15438627.2019.1660879

Sampaio, J., & Maças, V. (2012). Measuring tactical behaviour in football. *International Journal of Sports Medicine*, *33*(5), 395–401. https://doi.org/10.1055/s-0031-1301320

Sparrow, W. A., Donovan, E., Van Emmerik, R., & Barry, E. B. (1987). Using relative motion plots to measure changes in intra-limb and inter-limb coordination. *Journal of Motor Behavior*, *19*(1), 115–129. https://doi.org/10.1080/00222895.1987.10735403

Yiannakos, A., & Armatas, V. (2017). Evaluation of the goal scoring patterns in European Championship in Portugal 2004 ., *8668*. https://doi.org/10.1080/24748668.2006.11868366

Table 1. Tactical variables used separated by groups.

Groups	Variables	Values
Notational	Time of possession	absolute value
	Successful pass	frequency
	Successful pass last third	frequency
	Shots	frequency
	Goal	frequency
Space occupation	Offensive team's effective playing space	mean, CV, ApEn
	Defensive team's effective playing space	mean, CV, ApEn
	Offensive team's length	mean, CV, ApEn
	Defensive team's length	mean, CV, ApEn
	Offensive team's width	mean, CV, ApEn
	Defensive team's width	mean, CV, ApEn
	Distance between offensive team's centroid and target	mean, CV, ApEn
	Distance between defensive team's centroid and target	mean, CV, ApEn
	Distance between team's centroid	mean, CV, ApEn
	Centroid Progression	absolute value
Displacement synchronization	Offensive team's synchronization X-axis	% mean, CV
	Defensive team's synchronization X-axis	% mean, CV
	Offensive team's synchronization Y-axis	% mean, CV
	Defensive team's synchronization Y-axis	% mean, CV

Forty-one variables were classified into three groups; notational (five variables), space occupation (twenty-eight variables), displacement synchronization (eight variables). Notational variables represent the total occurrence of the offensive team's ball possession, except time of possession. All continuous space occupation variables are calculated as mean, coefficient of variation (CV), and approximate entropy (ApEn) per ball possession per each team, except centroid progression that represents the difference between offensive team's centroid position in the last ball possession moment and the beginning of ball possession. For all displacement synchronization variables mean values of the percentage (% mean) of all the dyads were calculated to represent the mean of team synchronization and the CV was calculated to indicate the variability between the dyads. Abbreviations: CV = coefficient of variation; ApEn = approximate entropy; % mean = mean of the percentage.

Table 2. Descriptive and inferential statistics of different clusters of ball possession sequences.

Variables	Short (Mean ± SD)	Medium (Mean ± SD)	Long (Mean ± SD)	Short vs Medium (Mean difference ± CL) Effect size	Short vs Long (Mean difference ± CL) Effect size	Medium vs Long (Mean difference ± CL) Effect size	F1 95.8%	F2 4.2%
Time possession	11.07^{ab} ± 4.49	26.83^{c} ± 7.33	55.50 ± 14.97	15.76 ± 1.07 very large	44.43 ± 2.09 very large	28.67 ± 2.96 very large	-	-
Successful pass	1.93^{ab} ± 0.99	5.41^{c} ± 1.84	12.11 ± 4.61	3.49; ± 0.26 very large	10.19; ± 0.59 very large	6.70; ± 0.84 very large	-	-
CV DEF-SynY	47.25^{ab} ± 13.55	30.11^{c} ± 8.10	20.89 ± 4.51	-17.14; ± 2,19 large	-26.37; ± 3.70 very large	-9.23; ± 2.29 large	,577*	.244
CV DEF-SynX	40.60^{ab} ± 14.43	27.00^{c} ± 8.44	20.77 ± 5.64	-13.60; ± 2.33 moderate	-19.84; ± 3.95 large	-6.24; ± 2.43 moderate	.418*	.246
CV OFF-DCT	9.39^{ab} ± 7.90	18.11^{c} ± 9.83	20.55 ± 8.99	8.72; ± 1.62 moderate	11.16; ± 2.37 large	2.44; ± 2.97 small	-.346*	-.343
mean OFF-WID	40.41^{ab} ± 6.84	45.97^{c} ± 6.13	49.21 ± 4.62	5.56; ± 1.22 moderate	8.80; ± 1.92 large	3.24; ± 1.80 small	-.335*	-.112
CV OFF-SynX	44.99^{ab} ± 16.03	35.10^{c} ± 11.30	27.81 ± 7.64	-9.88; ± 2.69 moderate	-17.18; ± 4.42 moderate	-7.30; ± 3.26 moderate	.289*	.014
CProgress	12.72^{b} ± 14.44^{a}	22.02^{c} ± 17.63	26.16 ± 13.07	9.30; ± 2.92 small	13.44; ± 4.18 moderate	4.13; ± 5.15 small	-.221*	-.136
mean OFF-DCT	55.00^{ab} ± 14.63	47.56 ± 9.48	44.97 ± 8.51	-7.44; ± 2.41 small	-10.03; ± 4.08 moderate	-2.59; ± 2.86 small	.210*	.174
mean DEF-LEN	34.34^{ab} ± 7.55	31.53^{c} ± 7.00	27.58 ± 6.85	-2.80; ± 1.37 small	-6.75; ± 2.19 moderate	-3.95; ± 2.15 small	.190*	-.174
% mean OFF-SynX	47.37^{ab} ± 15.85	42.52^{c} ± 9.84	36.56 ± 6.20	-4.85; ± 2.58 small	-10.81; ± 4.35 moderate	-5.96; ± 2.82 moderate	.172*	-.119
% mean DEF-SynX	47.84^{ab} ± 13.90	44.84 ± 10.47	42.85 ± 7.51	-3.00; ± 2.37 small	-4.99; ± 3.85 small	-4.03; ± 3.30 small	.097*	.018
Successful pass-LT	0.46^{ab} ± 0.83	1.58^{c} ± 1.68	3.38 ± 3.19	1.11; ± 0.22 moderate	2.92; ± 0.43 large	1.80; ± 0.65 moderate	-.381	.436*
% mean DEF-SynY	37.65 ± 11.62	37.16 ± 7.82	40.20 ± 6.57	-0.49; ± 1.92 trivial	2.54; ± 3.23 small	3.03; ± 2.33 small	-.025	.239*

Mean ± Standard deviation (SD), mean difference and respective 95% confidence limit (CL), effect size based on Cohen´s *d*, structure coefficient (SC) of 12 variables selected by the FDA model, and 2 variables used to separate the clusters (time of possession and successful pass). *variable better explained by function 1 or 2. One-way ANOVA and the Bonferroni post hoc to differentiate between groups (a = difference between clusters 1 and 2; b = difference between clusters 1 and 3; c = difference between clusters 2 and 3; $p<0.05$). Abbreviations: Short = Short ball possession sequences; Medium = Medium ball possession sequences; Long = Long ball possession sequences; F1 = Function 1; F2 = Function 2.

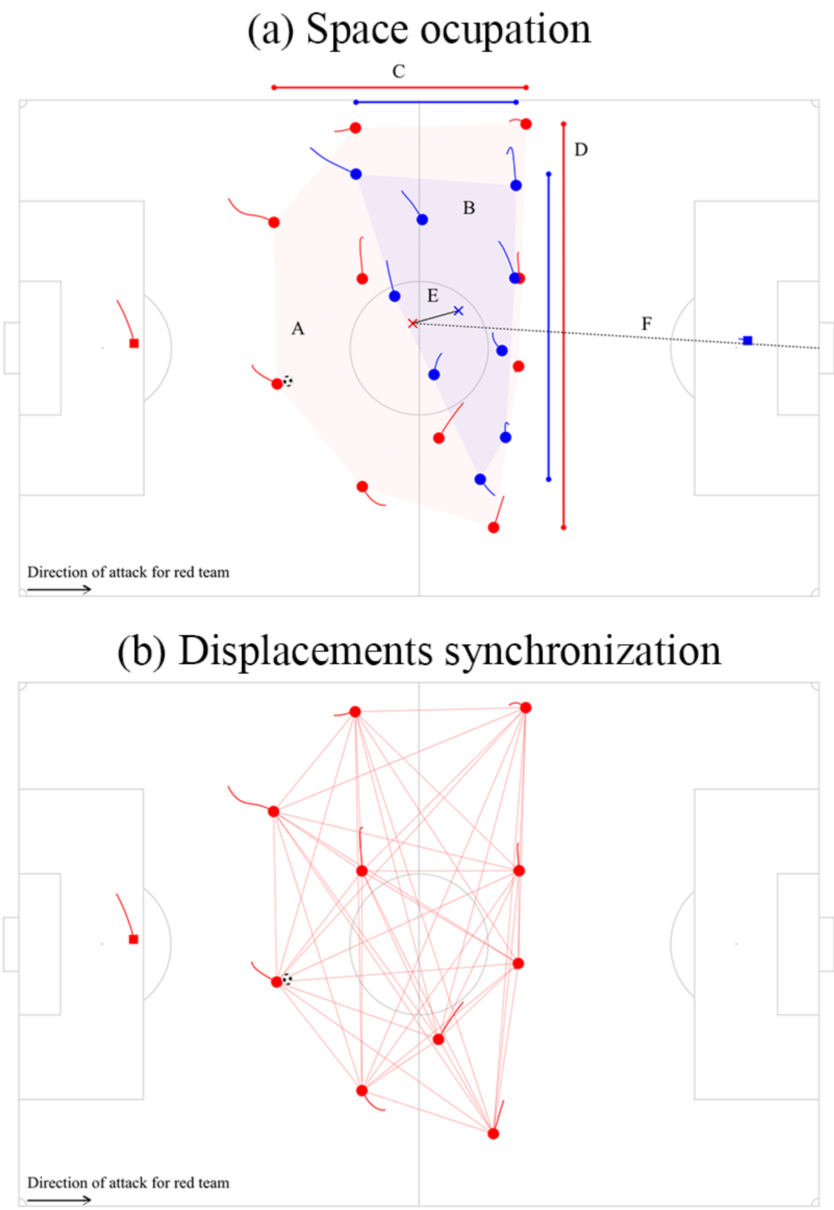

Figure 1.a) Representation of space occupation variables. Red team in ball possession (offensive phase) versus blue team (defensive phase) during long ball possession sequence. Abbreviations: A = Effective playing space (red team); B= Effective playing space (blue team); C = length (red team); D = width (red team); E = distance between team centroids; F = distance between centroid and target (red team). b) Representation of displacements synchronization. Each edge represents a dyad. Each player is connected to nine other players, except for the goalkeeper (total of 45 dyads).

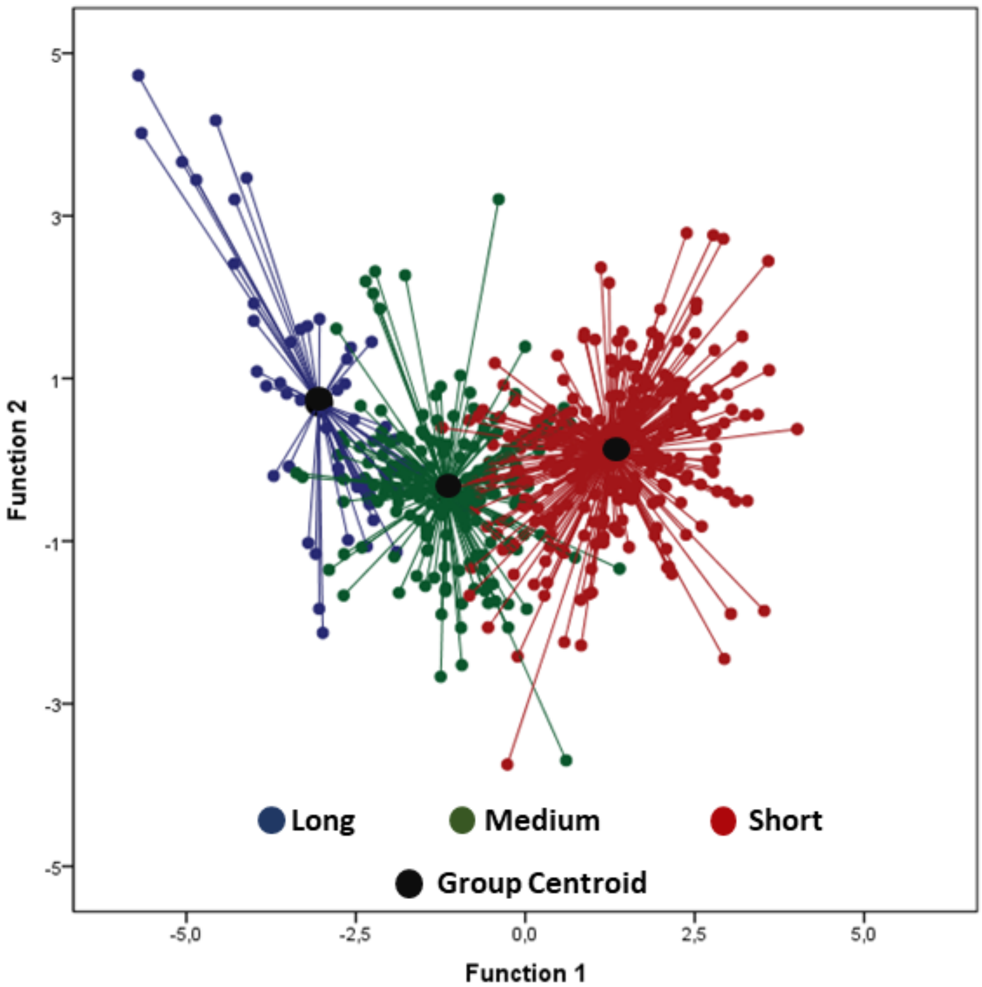

Figure 2. Territorial maps of the cluster centroids (group centroid) and their respective ball possession sequences (short = short ball possession; medium = medium ball possession; long = long ball possession) based on two canonical discriminant functions. Function 1 representing 95.8% of the total variance (0.83 of the canonical correlation) and function 2 representing 4.2% (0.30 of the canonical correlation), both functions being statistically significant ($p < 0.0001$), (Wilks' Lambda = 0.27 and 0.91 for functions 1 and 2, respectively).