SPRINGER BRIEFS IN
APPLIED SCIENCES AND TECHNOLOGY

Filipe Manuel Clemente
João Bernardo Sequeiros
Acácio F.P.P. Correia
Frutuoso G.M. Silva
Fernando Manuel Lourenço Martins

Computational Metrics for Soccer Analysis

Connecting the Dots

Springer

SpringerBriefs in Applied Sciences and Technology

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Systems Research Institute, Warsaw, Poland

SpringerBriefs present concise summaries of cutting-edge research and practical applications across a wide spectrum of fields. Featuring compact volumes of 50–125 pages, the series covers a range of content from professional to academic.

Typical publications can be:

- A timely report of state-of-the art methods
- An introduction to or a manual for the application of mathematical or computer techniques
- A bridge between new research results, as published in journal articles
- A snapshot of a hot or emerging topic
- An in-depth case study
- A presentation of core concepts that students must understand in order to make independent contributions

SpringerBriefs are characterized by fast, global electronic dissemination, standard publishing contracts, standardized manuscript preparation and formatting guidelines, and expedited production schedules.

On the one hand, **SpringerBriefs in Applied Sciences and Technology** are devoted to the publication of fundamentals and applications within the different classical engineering disciplines as well as in interdisciplinary fields that recently emerged between these areas. On the other hand, as the boundary separating fundamental research and applied technology is more and more dissolving, this series is particularly open to trans-disciplinary topics between fundamental science and engineering.

Indexed by EI-Compendex and Springerlink.

More information about this series at http://www.springer.com/series/8884

Filipe Manuel Clemente
João Bernardo Sequeiros
Acácio F.P.P. Correia · Frutuoso G.M. Silva
Fernando Manuel Lourenço Martins

Computational Metrics for Soccer Analysis

Connecting the Dots

Springer

Filipe Manuel Clemente
Escola Superior de Desporto e Lazer
Instituto Politécnico de Viana do Castelo
Melgaço
Portugal

and

Delegação da Covilhã
Instituto de Telecomunicações
Covilhã
Portugal

João Bernardo Sequeiros
Delegação da Covilhã
Instituto de Telecomunicações
Covilhã
Portugal

Acácio F.P.P. Correia
Delegação da Covilhã
Instituto de Telecomunicações
Covilhã
Portugal

Frutuoso G.M. Silva
Universidade da Beira Interior
Covilhã
Portugal

and

Delegação da Covilhã
Instituto de Telecomunicações
Covilhã
Portugal

Fernando Manuel Lourenço Martins
Delegação da Covilhã
Instituto de Telecomunicações
Covilhã
Portugal

and

Departamento de Educação, Escola Superior de Educação
Instituto Politécnico de Coimbra, IIA, RoboCorp, ASSERT
Coimbra
Portugal

ISSN 2191-530X ISSN 2191-5318 (electronic)
SpringerBriefs in Applied Sciences and Technology
ISBN 978-3-319-59028-8 ISBN 978-3-319-59029-5 (eBook)
https://doi.org/10.1007/978-3-319-59029-5

Library of Congress Control Number: 2017947026

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Acknowledgements

The authors would like to thank Instituto de Telecomunicações—Covilhã, Instituto Politécnico de Viana do Castelo—Escola Superior de Desporto e Lazer de Melgaço, Instituto Politécnico de Coimbra, Escola Superior de Educação, and Universidade da Beira Interior for the institutional support to make this book.

The authors would like also to thank to Abel João Padrão Gomes, Juan Pablo Sánchez, and Nuno Pinto for their scientific contribution in previous works and for their suggestions for this book.

Finally, the authors would like to thank to their families for the permanent support and for their patience with their scientific activity. For that reason, this book is dedicated to our families.

This work was carried out within the scope of R&D Unit 50008, financed by UID/EEA/50008/2013. This study was conducted in the aim of the granted project: uPATO from Instituto de Telecomunicações.

Contents

About the Authors

Filipe Manuel Clemente is Assistant Professor in Instituto Politécnico de Viana do Castelo, Escola Superior de Desporto e Lazer (Portugal) and researcher in Instituto de Telecomunicações, Delegação da Covilhã (Portugal). He has a Post. Doc. in social network analysis applied in team sports in Instituto de Telecomunicações (2016). He has a Ph.D. in Sport Sciences—Sports training in Faculty of Sport Sciences and Physical Education from University of Coimbra (2015, Portugal). His research in sports training and sports medicine has led to more than 130 publications, 43 of them with impact factor (indexed at JCR). He has conducted studies in performance analysis, match analysis, computational tactical metrics, network analysis applied to team sports analysis, small-sided and conditioned games, physical activity and health and sports medicine. He is currently guest editor of Science in Soccer special issue at Human Movement. He was guest editor on Sports Performance and Exercise Collection in Springer Plus journal (2015 and 2016) and Performance in Soccer special issue at Sports MDPI journal (2016). He is also editor in three more scientific journals and reviewer in nine journals indexed on Web of Knowledge.

For further details see: www.researchgate.net/profile/Filipe_Clemente

Contact: filipe.clemente5@gmail.com

João Bernardo Sequeiros is currently attending a Ph.D. in Computer Science and Engineering at Universidade da Beira Interior. He has a Master's Degree in Computer Science and Engineering at Universidade da Beira Interior, concluded in 2016, and a Bachelor's Degree in Computer Science and Engineering at Universidade da Beira Interior, concluded in 2014. His main research and interest areas are network and application security, Internet of Things, cryptography, and game development. He enjoys developing in C, Java, Python, and C# and likes to challenge his knowledge in network management, modeling software, and database management systems.

Contact: bernardo.seq@penhas.di.ubi.pt

Acácio F.P.P. Correia is a student who recently finished a Master's Degree in Computer Science and Engineering, at the Universidade da Beira Interior, his dissertation focused on the study of Natural Language Processing techniques and scientific document suggestions according to the context. He has previously concluded a Bachelor's Degree in Computer Science and Engineering, at the same university. His professional interests include procedural generation, artificial intelligence, natural language processing, cryptography, and video game development.

Contact: acaciofilipe.correia@gmail.com

Frutuoso G.M. Silva is Assistant Professor of the Department of Computer Science at the Universidade da Beira Interior (UBI) and leader of the Regain group of Instituto de Telecomunicações (IT). Currently, he is the coordinator of the Master's Degree in Game Design and Development of the UBI. His current research interests include geometric modeling, augmented reality, and computer games. He is a member of the Eurographics.

For further details see: http://www.di.ubi.pt/ fsilva/

Contact: fsilva@di.ubi.pt

Fernando Manuel Lourenço Martins is Research Member and Scientific Coordinator of Applied Mathematics group in Instituto de Telecomunicações, Delegação da Covilhã, Portugal and Professor and Course Director of Basic Education in Instituto Politécnico de Coimbra, Escola Superior de Educação, Department of Education (Portugal). He has a Ph.D. in Mathematics in Universidade da Beira Interior (Portugal). His research in applied mathematics and statistical analysis has led to more than 130 publications. The research included advances in network analysis applied to team sports and assessment of interactions between children, statistical analysis in team sports, and mathematical and statistical knowledge for teaching. He was co-editor on Sports Performance and Exercise Collection in Springer Plus Journal.

For further details see: https://www.it.pt/Members/Index/1877 and www.researchgate.net/profile/FernandoMartins13

Contact: fmlmartins@ubi.pt

Acronyms

DBMS	Database Management System
DHT	Discrete Hilbert Transform
GPS	Global Positioning System
LPM	Local Position Measurement
lpwratio	Length per Width Ratio
PTZ	Pan-Tilt-Zoom
RMSE	Root-Mean-Squared Error
SSG	Small-Sided Game
uPATO	Ultimate Performance Analysis Tool

Chapter 1
Brief Review About Computational Metrics Used in Team Sports

Abstract The purpose of this chapter is to analyze how position data have been used in the aim of match analysis. A brief related work will present the main measures and results that come from soccer analysis based on georeferencing. Individual measures that characterize the time-motion profile, tactical behavior, predictability, stability and spatial exploration of players will be discussed. Collective measures that represent the Geometrical Center and team's dispersion will be also presented during this chapter. The main evidences that resulted from these measures will be briefly discussed.

Keywords Position data · Georeferencing · Match analysis · Measures · Tactics · Soccer

1.1 Using the Dots to Characterize Individual Behavior: Related Work

Tracking systems have been used to follow the instantaneous positions of soccer players in both training and game situations [6]. These tracking systems determine the position of a player in a Cartesian coordinate system [35, 41], representing each player as a dot. Multicamera semi-automatic systems, local position measurement technology and global positioning system technology are the most used options to quickly record and process the data of players' positions on a pitch [3].

The wide use of these tracking systems provided new possibilities of analysis in the case of soccer. New applications and techniques are now used to determine and estimate individual and collective performance of players and teams [20]. Sports sciences are now capable of using the knowledge from scientific areas such as mathematics, physics, computer science and engineering to objectively measure performance, adjust training plans and make better decisions during a match [10].

The most common application of position data to support the daily practice of coaches and sport scientists is the calculus of time-motion profiles of players during training sessions and matches [5]. By using time-motion profiles, it is possible to

F.M. Clemente et al., *Computational Metrics for Soccer Analysis*, SpringerBriefs in Applied Sciences and Technology, https://doi.org/10.1007/978-3-319-59029-5_1

determinate the external load of players during their activities, namely measuring walking, jogging, running, striding and sprinting distances and characterizing the high-intensity activities as accelerations and decelerations [31]. Individualization of these data provides immediate and significant information regarding training loads and the impact on each player [33]. Moreover, motion-analysis also allows to determine the extent of fatigue experienced by players during competitions and identify the variance of performance during the season [6].

Position data have also been used to analyze the behavior of players [16, 39, 41]. Players and the ball were tracked in a single match as a case-study [41]. In that study, it was observed that all players, except the goalkeeper, follow the movements of the ball more closely in the longitudinal axis than in the lateral one [41]. However, in the same study it was found that the ball changed direction in the lateral axis with more frequency than in longitudinal one.

The variability of movements has also been analyzed through the use of Kolmogorov Entropy, Sample Entropy or Approximate Entropy [24]. Such approach allows for the identification of positional variability of a player in a time-series and to understand the regularity of their movements during the game [39]. The information of a position in a histogram can also be measured in its variability by using the Shannon Entropy [17].

In a single-match case study it was found that the higher values of Approximate Entropy and Shannon Entropy were found in midfielders [17]. In an alternative application, it was found that increasing the pitch size (as a task constraint) would result in a decrease of variability, thus the individual player zone was smaller [39].

The exploration of the pitch can be another interesting use of position data in soccer. In a study that analyzed the influence of pitch area-restrictions on tactical behavior, it was found that the Spatial Exploration Index can be higher under free-spacing conditions [29].

As possible to identify, position data provide a lot of possibilities to analyze the motion of players and characterize the behaviors in different scenarios. The use of such approaches may provide information to adjust Small Sided Games (SSGs) (smaller versions of soccer games) [14] and to add specific task constraints that may help in the re-adjustment of individual tactical behavior [19]. The use of position data to characterize individual behavior is growing and for that reason the major applications have not been published, so far.

However, not only individual behavior can be measured by using position data. The synchronization with an individual opponent (dyad), with teammates (cooperation) or with the opponents (opposition) is another possibility, used to objectively measure and determine patterns of interaction. Based on that, Sect. 1.2 will briefly present the related work about the use of position data to characterize the collective behavior of players during the games.

1.2 Measuring the Collective Behavior Based on Data-Position: Related Work

Interaction between teammates can be interpreted and measured in many different ways [20]. Traditional and new approaches to the match analysis provide new insight about the information that can be used to characterize the teams and their process of cooperation and opposition with other teams [40]. For that reason, the analysis can be done on an intra-level, to characterize the networking of a team, or on an inter-level, to identify the "rapport de forces" (balance of strengths) [30].

Collective organization can be understood as a dynamic process that emerges from the context based on the constraints of the match [34]. The interactions among teammates may be measured with a position data analysis [10]. In a spatio-temporal analysis we may identify the synchronization of teammates and the shapes of the team in different playing scenarios [2].

Some measures based on position data have been proposed to classify the collective behavior of teams and to identify the synchronization with the opponents [20, 40]. The classification of the measures can be based on their aim [1]: (i) team center; (ii) team dispersion; (iii) team synchronization; and (iv) division of labor within teams.

Measures that analyze the team center aim at identifying the Geometrical Center of a set of dots (in this case, the players of a team) [9]. Three different approaches have been conducted to compute the Geometrical Center of a team: (i) the centroid, that represents the exact Geometrical Center of the dots (players), excluding the goalkeeper [27]; (ii) the weighted centroid, which considers all players but requires the position of the ball to attribute weights to the position of each player in order to compute the centroid [7]; and (iii) the centroid considering the middle of the two farthest players [32]. The team center has been used to identify the variance of the middle point of a team during the match, to analyze the in-phase relationships between centroids of both teams and to identify the oscillations of centroids in critical moments (e.g., shots, goals) [2, 13, 27, 38].

The purpose of team dispersion is to objectively measure the expansion of a team and the areas covered by the players [28]. Usually, the dispersion of the teammates depends both on the match and ball possession statuses [18]. A higher dispersion of teammates can be observed in attacking moments than in defensive ones [36]. Different measures based on position data have been proposed: (i) Stretch Index, which can be described as the mean dispersion of the players, considering the Geometrical Center [21, 41]; (ii) weighted Stretch Index, which considers the dispersion of the players to the weighted centroid [7]; (iii) team's spread, which also considers the overall dispersion of all players [36]; (iv) Surface Area, which represents the area of a polygon constituted by all teammates [22, 27]; (v) effective area of play, which only considers the effective triangulations made by teammates [13]; (vi) playing area, which represents the area of maximum width and length of a team, in each instant [26]; (vii) team length and width, which measures the dispersion in both axes [23]; and (viii) defensive play area and triangulations, which measures the area covered by each sector of a team during defensive moments [12].

The main results of team's dispersion revealed that teams occupy a bigger area while in possession of the ball than in defensive moments [11]. Moreover, some studies found that bigger areas occupied in defensive moments are associated with critical moments such as suffered shots or goals [36]. The analysis of the synchronization about the covered areas of both teams revealed a tendency to an in-phase relationship, with the exception of short periods of anti-phase [37]. These exceptions of anti-phase were associated with critical moments (e.g., shots and goals) [37].

The analysis of variability of these covered areas also revealed a tendency to be more regular throughout the match, with a higher variability in the first minutes of the game [22]. The comparison with performance variables suggests that teams cover greater areas against weaker teams [4] and that areas are bigger in drawing matches than in losing or winning matches [11]. Finally, the triangulations made between midfielders occupy a greater area than those between the other playing roles [15].

Different measures based on position data have been proposed to identify the properties of teams and the behaviors of players. The major ranges have been used to classify the division of labor among players [41]. The main evidences suggest that defenders increase the individual covered area in attacking moments and that forwards increase the areas in defending moments [20]. Voronoi diagrams have also been used to characterize the spatial behavior of players and teams [25].

The axes "drawn" by the defensive, midfield and forward sectors were also analyzed by a specific measure [8]. The Sectorial Lines measure represents the line of a set of dots (players) in a specific region of a team. The synchronization of these three lines (defensive, midfield and forward) was small, thus suggesting an independent angular positioning between sectors during attacking and defensive moments [8].

As possible to identify, all these collective measures provide an interesting approach to the analysis of the match and may reveal some patterns of interaction between teammates and between teams, which can be determinant to optimize the training process and to make informed decisions about the strategy to adopt in a game.

The following sections will discuss the usability of these measures in match analysis and how to compute them in a dedicated software called Ultimate Performance Analysis Tool (uPATO).

References

1. Araújo D, Silva P, Davids K (2015) Capturing group tactical behaviors in expert team players. Handbook of sport expertise. Routledge, Abingdon, pp 209–220
2. Bartlett R, Button C, Robins M, Dutt-Mazumder A, Kennedy G (2012) Analysing team coordination patterns from player movement trajectories in football: methodological considerations. Int J Perform Anal Sport 12(2):398–424

3. Buchheit M, Allen A, Poon TK, Modonutti M, Gregson W, Di Salvo V (2014) Integrating different tracking systems in football: multiple camera semi-automatic system, local position measurement and gps technologies. J sports sci 32(20):1844–1857
4. 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 13:822–832
5. Christopher C (2013) Interpreting physical performance in professional soccer match-play: should we be more pragmatic in our approach? Sports Med 43(8):655–663
6. Christopher C, Bloomfield J, Nelsen L, Reilly T (2008) The role of motion analysis in elite soccer. Sports Med 38(10):839–862
7. Clemente FM, Couceiro MS, Martins FML, Mendes RS, Figueiredo AJ (2013) Measuring tactical behaviour using technological metrics: case study of a football game. Int J Sports Sci Coach 8(4):723–739
8. Clemente FM, Couceiro MS, Martins FML, Mendes RS, Figueiredo AJ (2014) Developing a football tactical metric to estimate the sectorial lines: a case study. Computational science and its applications. Springer, Berlin, pp 743–753
9. Clemente FM, Couceiro MS, Martins FML, Mendes RS, Figueiredo AJ (2014) Intelligent systems for analyzing soccer games: the weighted centroid. Ing e Investig 34(3):70–75
10. Clemente FM, Couceiro MS, Martins FML, Mendes RS, Figueiredo AJ (2014) Practical implementation of computational tactical metrics for the football game - towards an augmenting perception of coaches and sport analysts. In: Murgante B, Misra S, Rocha AMAC, Torre CM, Rocha JG, Falcão MI, Taniar D, Apduhan BO, Gervasi, O (eds) ICCSA (1). Lecture Notes in Computer Science, vol 8579. Springer, Berlin, pp 712–727
11. Clemente FM, Couceiro MS, Martins FML, Mendes RS, Figueiredo AJ (2014) Using collective metrics to inspect spatio-temporal relationships between football players. South African J Res Sport Phys Educ Recreat 36(2):47–59
12. Clemente FM, Couceiro MS, Martins FML, Mendes RS, Figueiredo AJ (2015) Soccer team's tactical behaviour: measuring territorial domain. Proc Inst Mech Eng Part P J Sports Eng Technol 229(1):58–66
13. Clemente FM, Couceiro MS, Martins FML, Mendes RS, Figueiredo AJ (2013) Measuring collective behaviour in football teams: inspecting the impact of each half of the match on ball possession. Int J Perform Anal Sport 13(3):678–689
14. Clemente FMC (2016) Small-sided and conditioned games in soccer training: the science and practical applications. Springer, Berlin
15. Clemente FMC, Silva F, Martins F, Kalamaras D, Mendes R (2016) Performance analysis tool (pato) for network analysis on team sports: a case study of fifa soccer world cup 2014. J Sports Eng Technol 230(3):158–170
16. Correia V, Araújo D, Fernandes O, Fonseca S, Davids K (2011) Territorial gain dynamics regulates success in attacking sub-phases of team sports. Psychol Sport Exerc 12:662–669
17. Couceiro MS, Clemente FM, Martins FML, Tenreiro Machado JA (2014) Dynamical stability and predictability of football players: the study of one match. Entropy 16(2):645–674
18. da Costa IT, Garganta J, Greco PJ, Mesquita I, Seabra A (2010) Influence of relative age effects and quality of tactical behaviour in the performance of youth soccer players. Int Perform Anal Sport 10:82–97
19. Davids K, Araújo D, Correia V, Vilar L (2013) How small-sided and conditioned games enhance acquisition of movement and decision-making skills. Exerc Sport Sci Rev 41(3):154–161
20. Duarte R, Araújo D, Correia V, Davids K (2012) Sports teams as superorganisms: implications of sociobiological models of behaviour for research and practice in team sports performance analysis. Sport Med 42(8):633–642
21. Duarte R, Travassos B, Araújo D, Richardson M (2013) The influence of manipulating the defensive playing method on collective synchrony of football teams. Perform Anal sport IX, 65
22. Duarte R, Araújo D, Folgado H, Esteves P, Marques P, Davids K (2013) Capturing complex, non-linear team behaviours during competitive football performance. J Syst Sci Complex 26(1):62–72

23. Folgado H, Lemmink KA, Frencken W, Sampaio J (2014) Length, width and centroid distance as easures of teams tactical performance in youth football. Eur J Sport Sci 14(sup1):S487–S492
24. Fonseca S, Milho J, Passos P, Araújo D, Davids K (2012) Approximate entropy normalized measures for analyzing social neurobiological systems. J Mot Behav 44(3):179–183
25. Fonseca S, Milho J, Travassos B, Araújo D, Lopes A (2013) Measuring spatial interaction behavior in team sports using superimposed voronoi diagrams. Int J Perform Anal Sport 13(1):179–189
26. Fradua L, Zubillaga A, Caro O, Iván Fernández-García A, Ruiz-Ruiz C, Tenga A (2013) Designing small-sided games for training tactical aspects in soccer: extrapolating pitch sizes from full-size professional matches. J Sports Sci 31(6):573–581
27. Frencken W, Lemmink K, Delleman N, Visscher C (2011) Oscillations of centroid position and surface area of football teams in small-sided games. Eur J Sport Sci 11(4):215–223
28. Frencken W, Van Der Plaats J, Visscher C, Lemmink KA (2013) Size matters: pitch dimensions constrain interactive team behaviour in soccer. J Syst Sci Complex 26(1):85–93
29. Gonçalves B, Esteves P, Folgado H, Ric A, Torrents C, Sampaio J (2017) Effects of pitch area-restrictions on tactical behavior, physical and physiological performances in soccer large-sided games. J Strength Cond Res. ahead-of-p
30. Gréhaigne JF, Godbout P, Zerai Z (2011) How the ‘rapport de forces’ evolves in a football match: the dynamics of collective decisions in a complex system. Rev Psicol del Deport 20(2):747–765
31. Halson SL (2014) Monitoring training load to understand fatigue in athletes. Sport Med 44(2):139–147
32. Lames M, Ertmer J, Walter F (2010) Oscillations in football-order and disorder in spatial interactions between the two teams. Int J Sport Psychol 41(4):85
33. Lovell R, Abt G (2013) Individualization of time-motion analysis: a case-cohort example. Int J Sport Physiol Perform 8(4):456–458
34. McGarry T (2005) Soccer as a dynamical system: some theoretical considerations. Proceedings of Science and Football. Taylor & Francis Group, London, pp 570–579
35. Memmert D, Lemmink KA, Sampaio J (2017) Current approaches to tactical performance analyses in soccer using position data. Sport Med 47(1):1–10
36. Moura FA, Martins LEB, De Anido RO, De Barros RML, Cunha SA (2012) Quantitative analysis of Brazilian football players’ organisation on the pitch. Sport Biomech 11(1):85–96
37. Moura FA, van Emmerik REA, Santana JE, Martins LEB, de Barros RML, Cunha SA (2016) Coordination analysis of players’ distribution in football using cross-correlation and vector coding techniques. J Sport Sci 34(24):2224–2232
38. Siegle M, Lames M (2013) Modeling soccer by means of relative phase. J Syst Sci Complex 26(1):14–20
39. Silva P, Aguiar P, Duarte R, Davids K, Araújo D, Garganta J (2014) Effects of pitch size and skill level on tactical behaviours of association football players during small-sided and conditioned games. Int J Sport Sci Coach 9(5):993–1006. Online date: Monday, December 22, 2014
40. Travassos B, Davids K, Araújo D, Esteves PT (2013) Performance analysis in team sports: advances from an ecological dynamics approach. Int J Perform Anal Sport 13(1):83–95
41. Yue Z, Broich H, Seifriz F, Mester J (2008) Mathematical analysis of a football game. part i: individual and collective behaviors. Stud Appl Math 121(3):223–243

Chapter 2
How to Use the Dots to Analyze the Behavior and the Collective Organization

Abstract Position data can be obtained from different approaches and techniques. Based on that, the aim of this chapter is to briefly present the main methods that have been used to track players during the games and training sessions and to discuss which kind of information can be used to analyze the individual and collective behavior. The introduction to the uPATO software will be also executed. This software allows to import position data from multi-camera tracking systems or Global Positioning System (GPS) and to compute tactical and collective measures representing group behavior.

Keywords Position data · Georeferencing · Tracking systems · GPS · uPATO · Soccer

2.1 Following the Players? Tracking Systems to Determine the Data Position

Tracking the movement of players during a match has become a common practice in sports [6]. The use of position data of players allows coaches to better understand the performance of players and of the team, giving them a competitive advantage over other teams. From position data, it is possible to calculate different metrics, both individual and collective, that give a better comprehension on training exercises and real game situations performance. From these measurements, match dynamics and team dependencies and coordination (both intra-team and inter-team) can be analyzed, as well as tactical patterns and their changes with regard to game moments, giving coaches and teams a powerful tool for performance analysis and data for metrics calculation [11].

Different tracking mechanisms and systems exist, each with their advantages and disadvantages. From multiple camera systems to local position systems or GPS systems, it is possible to track the position of players with good accuracy, and extract position data on every measured instant across a defined time interval. From the data, it is possible to calculate metrics and extrapolate information on the players and the team, which allows for a more comprehensive analysis of a game [12].

F.M. Clemente et al., *Computational Metrics for Soccer Analysis*, SpringerBriefs in Applied Sciences and Technology, https://doi.org/10.1007/978-3-319-59029-5_2

The following sections give some insight on different tracking systems for the position of players on set time instants.

2.1.1 Camera Systems

Camera systems record live footage of a game, which is then processed to detect each player on the field. Object detection can be made with different procedures, such as background subtraction or shape detection. From this, it is then possible to obtain location data of the players for that frame, which represents a time instant. For improving coverage of the field, some systems utilize multiple cameras, placed at different angles, where each camera covers an area of the field [12]. Another possibility is the use of rotating and zoom-enabled cameras (such as Pan-Tilt-Zoom (PTZ) cameras) [7, 8]. The full camera set captures the entire playing field. Captured frames are either fused first for detection, or first individually analyzed on each camera, and then the detected players on each camera are joined. From the detected players, tracking algorithms are used to compute the movement of the players across frames and, from there, calculate the positions of each player across the frames.

Camera systems provide some significant advantages over other tracking methods, mainly the possibility of tracking official games, since players are not allowed to have devices on them during these matches. Some systems, such as DatatraX and Feedbacksport, allow for cameras to be moved to different fields, but the majority of camera systems require a fixed installation, meaning that teams only have data on their home matches [6]. However, these systems also have some issues and disadvantages, mainly in terms of occlusion, change in light conditions, the need to install and calibrate cameras, the computational weight of the image processing and the time it takes to have all data registered [6].

2.1.2 GPS Systems

Global Positioning System is most commonly used in military or commercial applications for location tracking and orientation. It is a system composed of 27 satellites, orbiting the Earth. Time information is sent from the satellites to a GPS receiver, and through time comparison (of the clock on the satellites with the clock on the receiver) a position is triangulated. A minimum of four satellites is required to triangulate the position accurately [9, 10].

This system can also be adopted as a position data tracking device for players. Each player is accompanied by a GPS receiver, which tracks its position at every measured instant. With this, it is possible to measure performance in sports in less restricted environments, and analyze real game situations through the gathered data [10]. Some systems also add other sensors to the GPS receiver, to track physiological data, such as heart rate or hydration.

Previous GPS units had several disadvantages when compared to other architectures, due to high acquisition cost, lack of precision and low report frequency, which made it difficult to gather reliable speed and acceleration data. Nowadays, most GPS devices have improved on these previous points, but maintain other disadvantages, such as the size of the devices, battery life, the amount of data generated due to the better reporting rate, and the inability to use GPS indoors, making the system unusable in indoor sports. GPS remains heavily dependant on satellite connection, making even outdoors unreliable in certain low coverage situations.

GPS devices have an average reliability on reported distance, which is reduced with the increase of velocity or changes in direction. When measuring longer movements, in time and distance, the error in measurement stays below 5%, but when measuring sprints the error can be as high as 77% for a 10 m sprint on a 1 Hz GPS device (1 Hz signifies that one report is made each second) [4].

GPS devices have other advantages though. They are not affected by lighting conditions, and can track players even when grouped up. They are also portable, not being restricted to a single field, and require practically no setup when compared with camera-based systems.

GPS tracking was the used method for the position data of players used on the uPATO framework, with recourse to GPS units (10 Hz, Accelerometer 1 kHz, FieldWiz, Paudex, Switzerland).

2.1.3 Other Systems

Other types of tracking systems exist, albeit not as popular as the ones presented in the previous sections. The first methods used for capture were through the recording of the match, and then manually measuring and annotating positions and movements, and comparing with videotaped movements of players (e.g., sprinting) to provide some measures of calibration and comparison [6]. Other, more recent techniques, include, for example, the TrakPerformance (Sportstec, Warriewood, NSW, Australia) system, where a pen is used on a drawing tablet to register the movement of a single player on the field [6], or the Local Position Measurement (LPM) system (Inmotio Object tracking v2.6.9.545, 45 Hz, Amsterdam, the Netherlands), where players wear a vest transmitter and several base stations are placed to capture the signals of the vests [5]. The first system has the advantage of being highly portable, inexpensive and usable virtually on any field, but it can only track one player, requires continuous operation and is dependent on the skill of who is registering the movement. The second system has the advantage of being usable both indoors and outdoors, when compared with a traditional GPS, but requires installation of the base stations and is, therefore, not portable between fields.

2.2 Introducing the uPATO Software: From GPSs to Data Import

The uPATO is a tool developed with the objective of providing users with an easy way for analyzing the performance of players at a collective and individual level, in team sports. It is composed of three major modules: a GPS module for analyzing GPS data, a module for analyzing matrices containing data from passes and goals and a module providing representations for comparing metrics. The subject of this book focuses on position data, thus centering the discussion around this tool's GPS module and omitting the other two modules.

The GPS data is organized and stored in a database, which requires the installation of MySQL, creation of a database and configuration of a user with access to it. In order to access the database and its data, the application presents a login menu, requiring the introduction of the username and password previously configured. After the login menu, the user has access to all the functions of the module: introduce/import data, calculate metrics and process data, visualize representations of the metrics and game animations.

The first step is then importing data to the database for future processing and consulting. An overview of the importation process is described in Sect. 2.2.1. While Sect. 2.2.2 briefly introduces the process of calculating the available metrics in this tool and Sect. 2.2.3 discusses how the results are presented to the user and how they become available for exportation.

2.2.1 Importing Data

A database was designed with team sports in mind, being organized into the following main tables:

- Location—consists in a field or location where a game took place;
- Team—consists in a team of players;
- Game—a combination of two teams, a location and a date;
- Player—each of the players;
- Data—the location of a player at each instant during a game.

Before importing data into the system, the involving structure needs to be created, including the location, teams, game and player. Each of these entities is created in a separate menu, through the submission of a specific form such as the one shown in Fig. 2.1, used in the submission of a new field (or location). In this case, the user fills the form with a name and, either with the width and height of the field, or with the GPS coordinates of the four corners of the field. The order of the corners does not matter, as long as they are inserted following any circular direction.

Another example of a form is displayed in Fig. 2.2 for the creation of a new player. In this case the required data include the name, number, position and team.

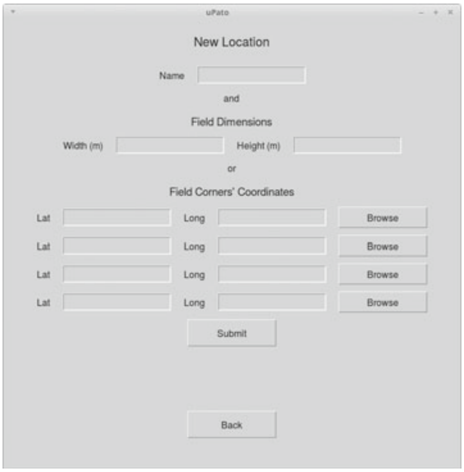

Fig. 2.1 Image of the form used in the creation of a new location

Fig. 2.2 Image of the form used in the creation of a new player

The system currently recognizes four different file formats of GPS data: the first two are similar, with the file describing the positions of a single player, where each row contains the instant of time and the position of that player (file formats obtained from the FieldWiz system [1]). The only difference between the two file formats is that one of them contains an initial matrix with the dimensions of the field before the rows with the actual data; the three other file formats include data from multiple players in a single file. One of them contains, in each row, the instant, position and name of the corresponding player (multiple players per file) and the other one describes files where each row contains an instant and the positions of a set of players in that instant (multiple players per row) (which is the file format used in the Johan system [2]).

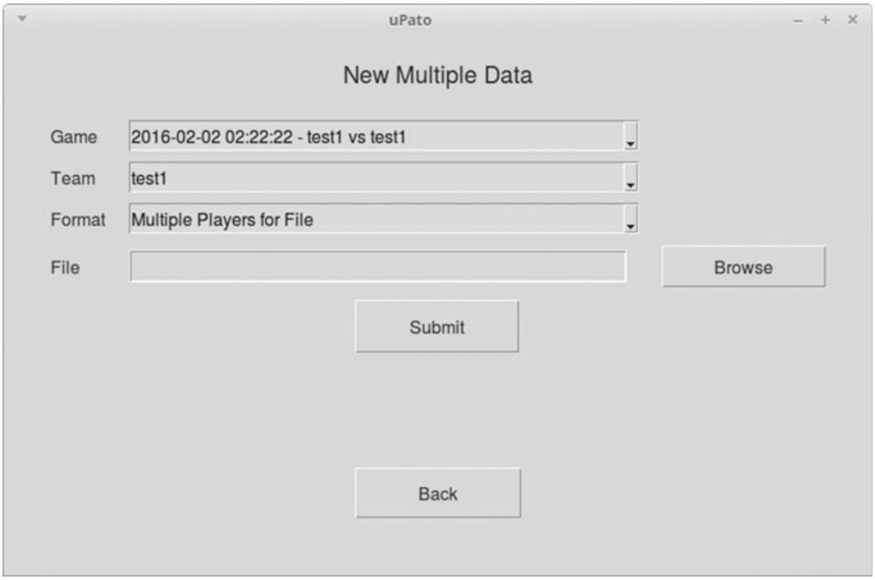

Fig. 2.3 Image of the form used in the importation of new data in one of the multiple players per file format recognized

Data from the TraXports system [3] are also recognized by the uPATO, even though its data isn't GPS based.

The introduction of data is divided into individual and multiple data formats, displaying a form for the selection of the game, player and specific file format in the first case and game, team and specific file format in the second. Fig. 2.3 presents the form used in the importation of data from a file from one of the multiple players per file format recognized, as previously described.

2.2.2 Processing Data

After inserting the data into the system, the user is able to calculate several metrics on the data. The uPATO divides metrics into two different sets: collective metrics, which involve data from all the players in a team; and individual metrics, referring to metrics calculated on the specific data of a player. A complete list of the metrics available for computing is presented in Appendix A.

In order to calculate the available metrics, two menus are available, one for each type, containing a specific form for selecting the data for processing, with the remaining steps being performed automatically by the system.

2.2.3 Results and Representations

When the calculation of the metrics is complete, representations are created with the results and presented to the user. The presented data is available for exportation at the press of a button, saving all data into .csv files and the representations into image files. With the data in .csv files, they can then be imported and analyzed in other systems or, in some cases, depending on the format, be passed to the data comparison module, which will then create barplots or boxplots of the data, allowing for a visual comparison and analysis.

Another function provided by uPATO is the creation of game animations with the position data of players. This option is available through the submission of a form selecting the data of interest, in the `Game Animation` menu. The game animation presents an animation of the selected period of time, where each player is represented by a dot positioned accordingly in the field and updated in real time. The animation also calculates most collective metrics, allowing users to visualize these metrics and some of their values at the same time as watching the game.

References

1. Fieldwiz (2017). http://www.fieldwiz.com/. Accessed 13 June 2017
2. Johan (2017). http://www.johan-sports.com/system/. Accessed 13 June 2017
3. Traxports (2017). http://www.ingeniarius.pt/?page=traxports. Accessed 13 June 2017
4. Aughey RJ (2011) Applications of GPS technologies to field sports. Int J Sports Physiol Perform 6(3):295–310
5. Buchheit M, Allen A, Poon TK, Modonutti M, Gregson W, Di Salvo V (2014) Integrating different tracking systems in football: multiple camera semi-automatic system, local position measurement and gps technologies. J Sports Sci 32(20):1844–1857
6. Carling C, Bloomfield J, Nelsen L, Reilly T (2008) The role of motion analysis in elite soccer. Sports Med 38(10):839–862
7. Du W, Hayet J-B, Piater J, Verly J (2006) Collaborative multi-camera tracking of athletes in team sports. Workshop on computer vision based analysis in sport environments (CVBASE) 5439:2–13
8. Hayet J-B, Mathes T, Czyz J, Piater J, Verly J, Macq B (2005) A modular multi-camera framework for team sports tracking. In: IEEE Conference on advanced video and signal based surveillance, AVSS 2005. IEEE, pp 493–498
9. Hofmann-Wellenhof B, Lichtenegger H, Collins J (1994) Overview of GPS. Global positioning system. Springer, Berlin, pp 1–11
10. Larsson P (2003) Global positioning system and sport-specific testing. Sports Med 33(15):1093–1101
11. Memmert D, Lemmink KAPM, Sampaio J (2017) Current approaches to tactical performance analyses in soccer using position data. Sports Med 47(1):1–10
12. Múller Junior B, Anido RDO (2004) Distributed real-time soccer tracking. In: Proceedings of the ACM 2nd international workshop on Video surveillance and sensor networks. ACM, pp 97–103

Chapter 3
Individual Metrics to Characterize the Players

Abstract The purpose of this chapter is to present the individual measures that can be computed in the uPATO software. Each measure will be presented with a definition and case-studies to discuss the data and how results can be interpreted. Time-motion profile (including distances at different speeds), Shannon Entropy, Longitudinal and Lateral Displacements to the goal and variability, Kolmogorov Entropy and Spatial Exploration Index will be presented and discussed in this chapter. The case studies presented involve two five-player teams in an SSG considering only the space of half pitch (68 m goal-to-goal and 52 m side-to-side) and another eleven-player team in a match considering the space of the entire field (106.744 m goal-to-goal and 66.611 m side-to-side) even though only playing in half pitch.

Keywords Position data · Georeferencing · uPATO · Soccer · Tactics · Time-motion

3.1 Time-Motion Profile

3.1.1 Basic Concepts

From the position data of a player, it is possible to compute both the distance covered between two measurements and the velocity at which that distance was covered. It is then possible to segment the covered distance in regard to the type of movement, based on the velocity of the player.

Definition 3.1 [19] Let P be a set of points where each point represents the position of a player on each measured time instant. The average velocity of a player is obtained through the following equation:

$$V = \frac{\sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}}{\Delta t}, \tag{3.1}$$

where (x_0, y_0) and (x_1, y_1) represent two consecutive positions of a player, and Δt represents the time interval between the measurements of the two points.

F.M. Clemente et al., *Computational Metrics for Soccer Analysis*, SpringerBriefs in Applied Sciences and Technology, https://doi.org/10.1007/978-3-319-59029-5_3

Definition 3.2 [17] The distance covered by a player with a given type of movement is calculated through the following equation:

$$Dist = \sum \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}, \tag{3.2}$$

where the average velocity calculated between (x_0, y_0) and (x_1, y_1) is within the threshold of the type of movement.

Remark 3.1 The separation of the distances covered by the player when walking, jogging, running and sprinting is defined by velocity thresholds. The chosen thresholds are defined in [13].

3.1.2 Real Life Examples

The results obtained by a player in the SSG are presented in Table 3.1 with intervals of 30 s and for the entire 3 min in Table 3.2.

These can be compared to those obtained by another player in the match, presented in Table 3.3 with intervals of 30 s and for the entire 3 min in Table 3.4.

Table 3.1 Values obtained for the Time-motion profile of a player in an SSG, for periods of 30 s

Period of time (s)	Walk dist (m)	Jog dist (m)	Run dist (m)	Sprint dist (m)	Total dist (m)
[0; 30[	14.0766	18.0841	7.5997	2.8521	42.6124
[30; 60[	19.6351	19.2887	25.9350	0	64.8589
[60; 90[	25.7007	12.2484	0	0	37.9491
[90; 120[	12.6366	41.3389	14.0171	0.6810	68.6736
[120; 150[	25.4634	0	0	0	25.4634
[150; 180]	15.7107	19.8147	7.8640	0	43.3894

Table 3.2 Values obtained for the Time-motion profile of a player in an SSG, in the entire period of time of 3 min

Period of time (s)	Walk dist (m)	Jog dist (m)	Run dist (m)	Sprint dist (m)	Total dist (m)
[0; 180]	113.2232	110.7749	55.4157	3.5331	282.9468

Table 3.3 Values obtained for the Time-motion profile of a player in a match, for periods of 30 s

Period of time (s)	Walk dist (m)	Jog dist (m)	Run dist (m)	Sprint dist (m)	Total dist (m)
[0; 30[	11.4048	21.8184	11.3642	0	44.5873
[30; 60[	16.2274	0	0	0	16.2274
[60; 90[	12.7803	15.3015	2.0716	0	30.1534
[90; 120[	14.6113	10.4555	0	0	25.0668
[120; 150[	19.8129	21.9296	8.9573	0	50.6997
[150; 180]	15.9083	32.2080	0	0	48.1163

Table 3.4 Values obtained for the Time-motion profile of a player in a match, in the entire period of time of 3 min

Period of time (s)	Walk dist (m)	Jog dist (m)	Run dist (m)	Sprint dist (m)	Total dist (m)
[0; 180]	90.7450	101.7130	22.3931	0	214.8510

3.1.3 General Interpretation

Time-motion analysis can be used to characterize the activity profile of players during matches and training sessions [5]. Typically, these analyses measure the distances made at different speeds and acceleration thresholds [2]. However, the customizability of speed thresholds afforded by GPS software resulted in a wide range of different speed zones and thresholds, thus being difficult to demarcate and compare different locomotor activities between studies [8]. A recent brief review proposed the following speed thresholds, in km h^{-1}, for moderate-, high-, very-high-speed running and sprint [15]: [0; 5[, [5; 10[, [10; 15[, [15; 20[and [25; $+\infty$[, respectively. In the case of uPATO software [6] the thresholds, in km h^{-1}, are: [0; 7[(walking); [7; 14[(jogging); [14; 20[(running); and [20; $+\infty$[(sprinting).

The use of absolute measures (total distance (m) in each speed threshold) or relative measures (e.g., m $\min^{-1}$) may also be important in the moment of interpretation. Relative measures may provide a more accurate reflection of match intensity than total distance covered, which only provides information about the volume [8]. This option of pace (m $\min^{-1}$) can be visualized in the uPATO software.

The use of Time-motion analysis may help in characterizing the external load of training sessions and matches on players and facilitate coach decision making [4]. The analysis can be made by absolute or cohort-specific speed zones (player-independent) or individualizing the thresholds by player according to his fitness level [1]. The use of individual fitness levels make possible to determine the dose response in competition situations and adjust the analysis to the individualized evolution during the season [15].

3.2 Shannon Entropy

3.2.1 Basic Concepts

Shannon Entropy is calculated based on a player's position. The entropy represents how static or dynamic a player's positioning on the field is. Higher values are related to a greater dispersion of the player on the field (e.g., midfielders present higher entropy values) while lower values represent a player with a more fixed position on the field (e.g., a goalkeeper).

Definition 3.3 [7] Let p_i be the probability mass function. Shannon Entropy is given by the following formula:

$$E = -\sum_i p_i log_2 p_i. \tag{3.3}$$

Definition 3.4 [7] The probability mass function is given by:

$$p_i = \frac{h_i}{N_c}, \tag{3.4}$$

where h_i represents the histogram entry of intensity value i and N_c is the number of total cells of the field (in this implementation, the number of cells is given by the area of the field, in m^2).

3.2.2 Real Life Examples

The results obtained by three players in the SSG are presented in Table 3.5 with intervals of 30 s and represented in Fig. 3.1, with the entire 3 min in Table 3.6 and the heatmap of player 3 in Fig. 3.2.

Table 3.5 Values obtained for the Shannon Entropy in an SSG, for periods of 30 s

Average point (s)	Period of time (s)	Player 1	Player 2	Player 3
15	[0; 30[	0.6935	0.6880	0.6591
45	[30; 60[	0.7858	0.8405	0.8778
75	[60; 90[	0.7248	0.6973	0.6953
105	[90; 120[	0.7718	0.8560	0.8289
135	[120; 150[	0.6573	0.6534	0.6243
165	[150; 180]	0.6556	0.7479	0.7077

Table 3.6 Values obtained for the Shannon Entropy in an SSG, in the entire period of time of 3 min

Period of time (s)	Player 1	Player 2	Player 3
[0; 180]	3.9059	4.2571	4.1987

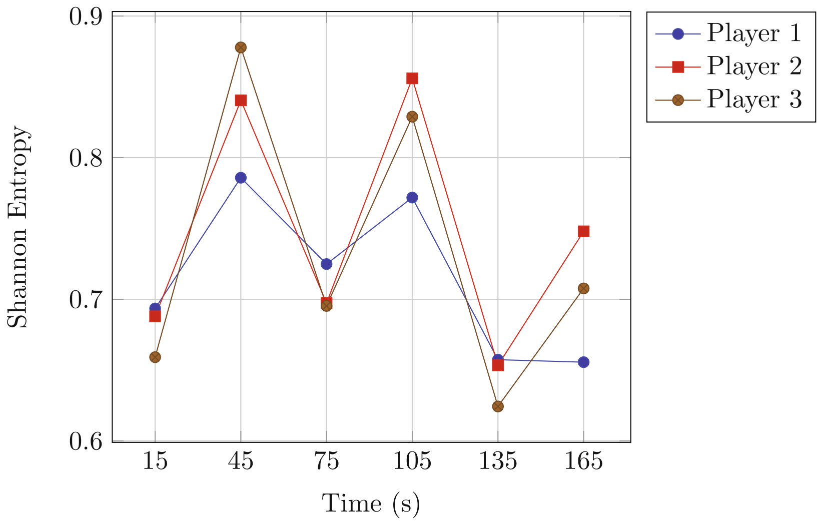

Fig. 3.1 Plotting of the values from Table 3.5, representing the Shannon Entropy in an SSG

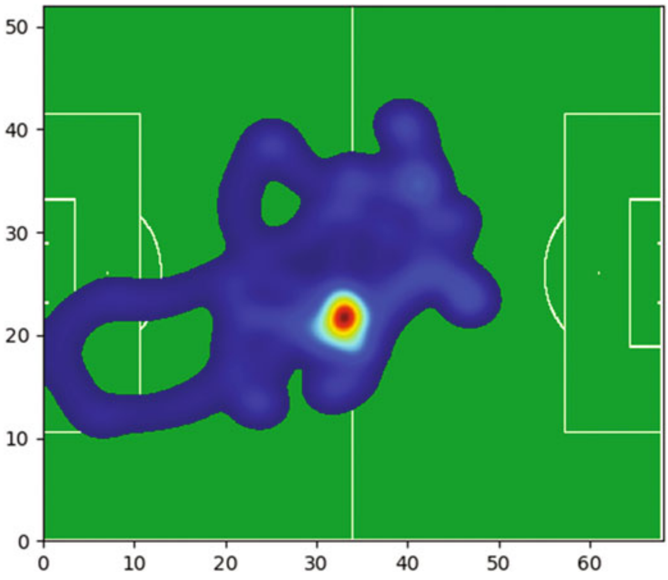

Fig. 3.2 Heatmap from player 3 in the SSG inside a half pitch of size 68 m goal-to-goal and 52 m side-to-side

These values can be compared to those obtained by another three players in the match, presented in Table 3.7 with intervals of 30 s and represented in Fig. 3.3, with the results for the entire 3 min in Table 3.8 and the heatmap from player 1 in Fig. 3.4.

Table 3.7 Values obtained for the Shannon Entropy in a match, for periods of 30 s

Average point (s)	Period of time (s)	Player 1	Player 2	Player 3
15	[0; 30[	0.3768	0.4071	0.3784
45	[30; 60[	0.3202	0.3498	0.4096
75	[60; 90[	0.3564	0.4160	0.4349
105	[90; 120[	0.3359	0.3768	0.3882
135	[120; 150[	0.4185	0.4385	0.4501
165	[150; 180]	0.3951	0.3560	0.3245

Table 3.8 Values obtained for the Shannon Entropy in a match, in the entire period of time of 3 min

Period of time (s)	Player 1	Player 2	Player 3
[0; 180]	2.0926	2.3032	2.3393

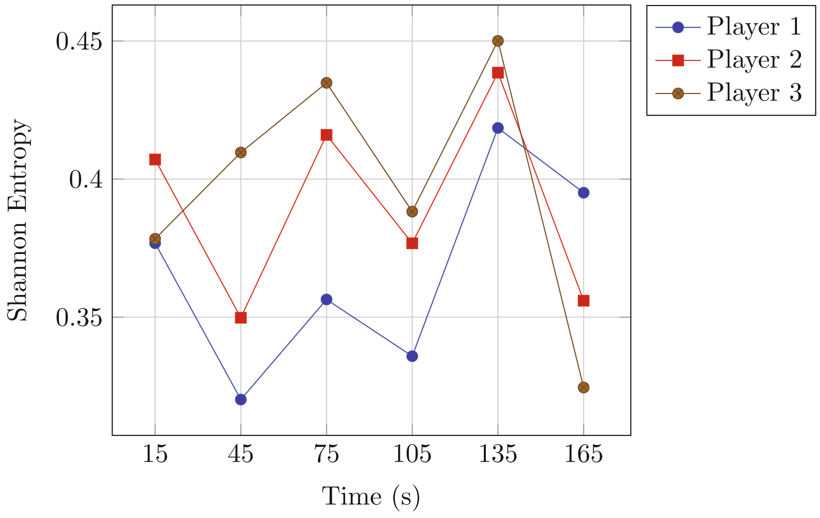

Fig. 3.3 Plotting of the values from Table 3.7, representing the Shannon Entropy in a match

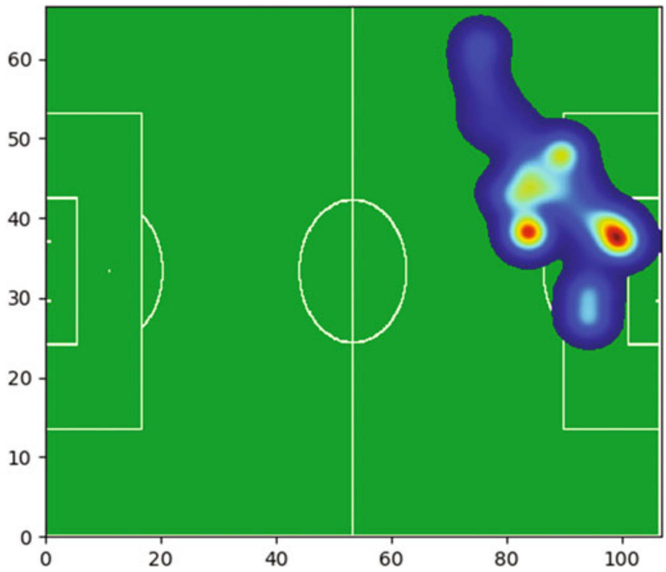

Fig. 3.4 Heatmap from player 1 in the match in a 106.744 m goal-to-goal and 66.611 m side-to-side field, while players only play in half pitch

3.2.3 General Interpretation

Heat maps represent the spatial distribution of a player considering the time spent in a specific position, thus being the frequency distribution of a player in the soccer pitch [7]. The Shannon Entropy can be used to classify heat maps, thus providing information about the spatial variability of the players in the pitch [7].

A value near 0 (zero) of Shannon Entropy suggests that the distribution is restricted and that the position of the player can be easily predicted [18]. On the other hand, higher values of Shannon Entropy indicate a more homogeneous distribution on the soccer pitch (uniform distribution), thus suggesting a high variability and that player is more unpredictable [18].

This measure can be used to identify the influence of different tasks during training sessions in the variability of the players. Moreover, the Shannon Entropy can measure the variability of position of players during the matches and classify the predictability of these players on their playing roles. Coaches may use this information, about the variability, to organize specific tasks that contribute to regulate the positioning behavior of players and to monitor the individualized evolution during a season.

3.3 Longitudinal and Lateral Displacements to the Goal and Variability

3.3.1 Basic Concepts

From the movement of a player across a set time interval, the angle between the player and the perpendicular line to the middle point of the goal can be calculated for every instant on the time interval, giving the angle to that line for the player. From the movement of the player it is also possible to calculate the displacement trajectory of the player in relation to the goal after a set time interval, these angles are best explained in Fig. 3.5.

Definition 3.5 [3, 10] Given a time-series of N instants, where each sample forms a pair of position coordinates, (x, y), and given a time interval of displacement measurements, T, the displacement angle of the movement of the player during the time interval T is given by:

$$A = arccos\left(\frac{\overrightarrow{p_1} . \overrightarrow{v_g}}{\sqrt{(p_{1_x}^2 + p_{1_y}^2) \times (v_{g_x}^2 + v_{g_y}^2)}}\right), \tag{3.5}$$

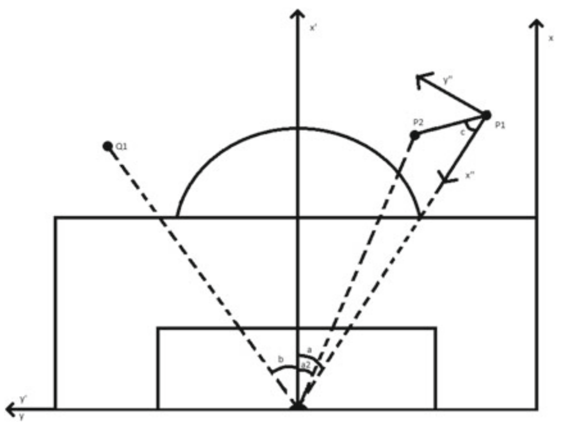

Fig. 3.5 Scheme for displacement angle and angle to goal calculation. The angle to goal of positions P_1, P_2 and Q_1 is given by the angles a, a_2 and b, respectively. These angles are calculated taking into consideration the axes x' and y'. This makes the angles to goal vary between 90° and 0° for player to the left of the perpendicular line to the goal, and between 0° and −90° for players located to the right. As an example, a and $a2$ will have values in the 0° to −90° range, while b will have a value in the 0° to 90° range. The displacement angle between two position of a player, P_1 and P_2, is represented as angle c. The axes taken into consideration on the calculation are axes x'' and y''. This angle can vary between −180° and 180°

where p_1 is the vector defined between the position of the player at the beggining and end of the time interval T, and v_g is the vector defined by the starting position of the player and the middle point of the goal.

Definition 3.6 [3, 10] Given a time-series of N instants, the angle, A, between the line defined by the player position on an instant i and the middle point of the goal and the perpendicular line to the middle point of the goal is calculated as follows:

$$A = arctan(m_{pos}), \tag{3.6}$$

where m_pos is the slope of the line defined by the player position on instant i and the middle point of the goal. The slope is obtained by the following equation:

$$m_{pos} = \frac{y_2 - y_1}{x_2 - x_1}, \tag{3.7}$$

where (x_1, y_1) is the position of the player on instant i and (x_2, y_2) is the position of the middle point of the goal.

3.3.2 Real Life Examples

The results obtained by three players in the SSG for the Longitudinal and Lateral Displacements are presented in Table 3.9 with intervals of 10 s and Angle to Goal in Table 3.10 with intervals of 0.1 s.

These values can be compared to those obtained by another three players in the match, presented in Table 3.11 with intervals of 10 s and the Angle to Goal values to those in Table 3.12 with intervals of 0.1 s.

Table 3.9 Values obtained for the Longitudinal and Lateral Displacements in an SSG, for periods of 10 s

Average point (s)	Period of time (s)	Player 1 (°)	Player 2 (°)	Player 3 (°)
5	[0; 10[	63.5019	53.9935	41.9927
15	[10; 20[	63.7860	53.3250	41.4363
25	[20; 30[	105.2644	63.8233	51.5511
35	[30; 40[	49.7925	43.2353	37.3341
45	[40; 50[	167.6322	38.6023	59.1726
55	[50; 60[	91.8993	71.6322	95.0328

Table 3.10 Values obtained for the angle to goal in an SSG, for periods of 0.1 s

Average point (s)	Period of time (s)	Player 1 (°)	Player 2 (°)	Player 3 (°)
0.05	[0; 0.1[	−11.5499	−13.4285	−30.6860
0.15	[0.1; 0.2[	−11.5499	−13.3934	−30.6911
0.25	[0.2; 0.3[	−11.5233	−13.3233	−30.6970
0.35	[0.3; 0.4[	−11.5147	−13.2494	−30.6970
0.45	[0.4; 0.5[	−11.4924	−13.2057	−30.7028
0.55	[0.5; 0.6[	−11.4572	−13.1813	−30.7129

Table 3.11 Values obtained for the Longitudinal and Lateral Displacements in a match, for periods of 10 s

Average point (s)	Period of time (s)	Player 1 (°)	Player 2 (°)	Player 3 (°)
5	[0; 10[	31.7342	36.9807	37.9728
15	[10; 20[	31.4339	36.6024	37.3645
25	[20; 30[	29.9888	33.7912	35.4670
35	[30; 40[	33.7496	40.0150	35.7440
45	[40; 50[	35.1627	39.4408	36.7123
55	[50; 60[	34.8835	41.3330	41.6626

Table 3.12 Values obtained for the angle to goal in a match, for periods of 0.1 s

Average point (s)	Period of time (s)	Player 1 (°)	Player 2 (°)	Player 3 (°)
0.05	[0; 0.1[	−3.7766	−10.5700	−0.7246
0.15	[0.1; 0.2[	−3.7713	−10.6240	−0.8028
0.25	[0.2; 0.3[	−3.7660	−10.6598	−0.8644
0.35	[0.3; 0.4[	−3.7660	−10.6897	−0.9055
0.45	[0.4; 0.5[	−3.7652	−10.7134	−0.9411
0.55	[0.5; 0.6[	−3.7637	−10.7491	−0.9917

3.3.3 *General Interpretation*

The player's position can be associated with the opponent's goal based on his position. This measure considers the variation in the pitch and allows coaches to understand the instantaneous variability of player. The measure will be useful for specific analysis to the shots or goals. Long-term analysis cannot provide real information for coaches. Longitudinal (goal-to-goal) and Lateral Displacement (side-to-side) trajectories of players can be associated with the middle point of goal line [12]. The distance of the players to the goal is measured and the displacement can be tracked instantaneously by considering the referencing point in the middle of the goal line.

This measure can be used to identify the variability of displacements made by players in specific critical moments (e.g., attacking building, defensive organization). Moreover, it can also be used to classify the collective synchronization in Longitudinal and Lateral Displacements during critical moments. The capacity to move simultaneously based on teammates' displacements can be determinant in specific moments, trying to ensure a cohesive structure and the unit [9].

3.4 Kolmogorov Entropy

The variability of a player across a game can be estimated through the Kolmogorov Entropy, giving a comparison mean between players in terms of positioning on the field, with higher values being associated with players that occupy larger areas of the field in the match.

3.4.1 Basic Concepts

Definition 3.7 [7] Given a time-series of N instants, represented as $u(1)$, $u(2)$, …, $u(N)$, where each sample forms a sequence of vectors $x(1)$, $x(2)$, …, $x(N-m+1) \in \mathbb{R}^{1\times m}$, each defined by the array $x(i) = [u(i)u(i+1)\cdots u(i+m-1)] \in \mathbb{R}^{1\times m}$, $C_i^m(\varepsilon)$ can be given by:

$$C_i^m(\varepsilon) = \frac{N_j}{N-m+1}, \tag{3.8}$$

where $N_j = \{\text{number of } x(j) \text{ such that } d(x(i), x(j)) \leq \varepsilon\}$ and the distance between $x(i)$ and $x(j)$ is given by:

$$d(x(i), x(j)) = \max_{k=1,2,\ldots,m} |u(i+k-1) - u(j+k-1)|. \tag{3.9}$$

Definition 3.8 [7] The Kolmogorov Entropy can be defined as:

$$KE = \Phi^m(\varepsilon) - \Phi^{m+1}(\varepsilon), \tag{3.10}$$

where $\Phi^m(\varepsilon)$ is given by:

$$\Phi^m(\varepsilon) = (N-m+1)^{-1} \sum_{i=1}^{N-m+1} ln\ C_i^m(\varepsilon). \tag{3.11}$$

3.4.2 Real Life Examples

The results obtained by three players in the SSG are presented in Table 3.13 with intervals of 30 s and represented in Fig. 3.6, with the results for the entire 3 min in Table 3.14.

These values can be compared to those obtained by another three players in the match, presented in Table 3.15 with intervals of 30 s and represented in Fig. 3.7, with the results for the entire 3 min in Table 3.16.

Table 3.13 Values obtained for the Kolmogorov Entropy in the SSG, for periods of 30 s

Average point (s)	Period of time (s)	Player 1	Player 2	Player 3
15	[0; 30[	0.2936	0.1936	0.1729
45	[30; 60[	0.3173	0.3778	0.3841
75	[60; 90[	0.1621	0.2247	0.0968
105	[90; 120[	0.4253	0.4308	0.2943
135	[120; 150[	0.0895	0.1942	0.1383
165	[150; 180]	0.3063	0.3509	0.2307

Table 3.14 Values obtained for the Kolmogorov Entropy in an SSG, in the entire period of time of 3 min

Period of time (s)	Player 1	Player 2	Player 3
[0; 180]	0.2840	0.3166	0.2455

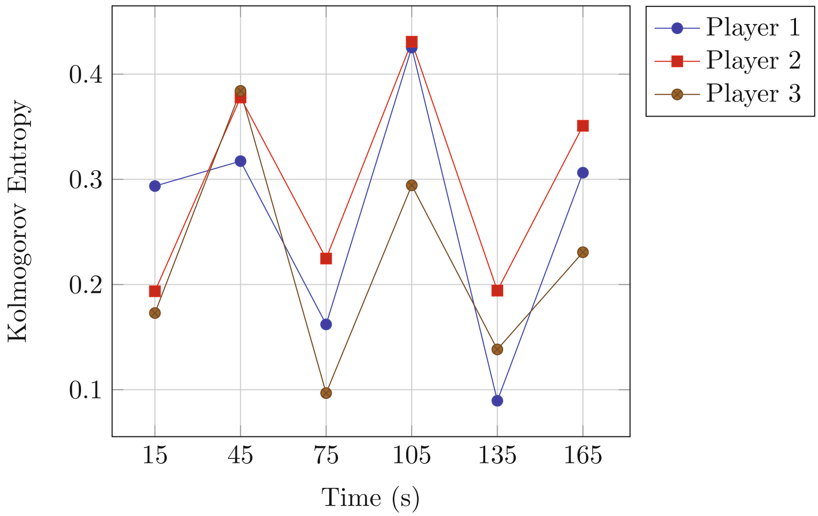

Fig. 3.6 Plotting of the values from Table 3.13, representing the Kolmogorov Entropy in an SSG

Table 3.15 Values obtained for the Kolmogorov Entropy in a match, for periods of 30 s

Average point (s)	Period of time (s)	Player 1	Player 2	Player 3
15	[0; 30[	0.1647	0.2113	0.1838
45	[30; 60[	0.0337	0.0932	0.1119
75	[60; 90[	0.1578	0.2068	0.2908
105	[90; 120[	0.1115	0.1737	0.1810
135	[120; 150[	0.2079	0.2972	0.2122
165	[150; 180]	0.1770	0.1284	0.1065

Table 3.16 Values obtained for the Kolmogorov Entropy in a match, in the entire period of time of 3 min

Period of time (s)	Player 1	Player 2	Player 3
[0; 180]	0.1540	0.2093	0.1989

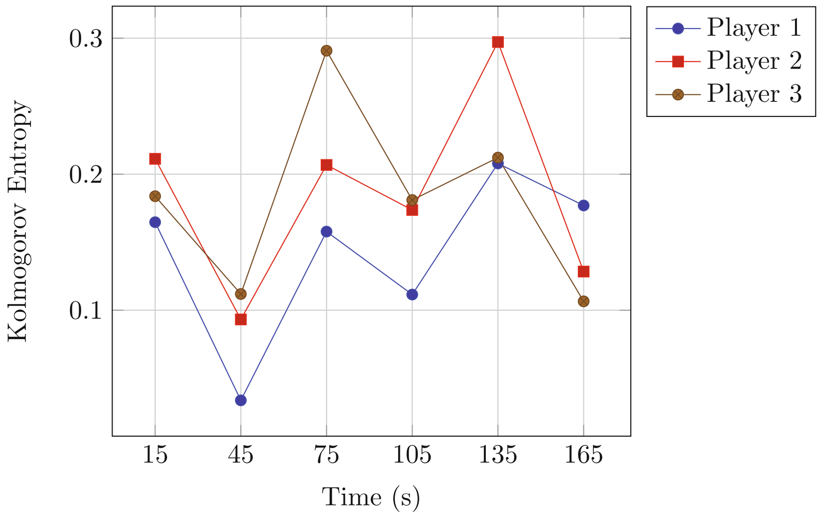

Fig. 3.7 Plotting of the values from Table 3.15, representing the Kolmogorov Entropy in a match

3.4.3 General Interpretation

The positioning variability considering a given time-series can be estimated by using the Kolmogorov Entropy [7]. The variation in the players' trajectories during the periods of the game can be used to classify the regularity of players in keeping a position or in occupying certain areas of play.

The Kolmogorov Entropy can be also used to analyze the variability of a set of collective measures such as center of the game, Stretch Index or Surface Area [11].

This entropy measure belongs to the non-linear statistics and can provide information about the regularity of individual and collective values.

Three main thresholds must be considered in the data interpretation of Kolmogorov Entropy [16]: $\tilde{0}$ (periodic function); 0.1 (chaotic system); and 1.5 random time series.

3.5 Spatial Exploration Index

3.5.1 Basic Concepts

From the position data, by measuring the average difference between a player's average position and its actual position on each measured instant, the Spatial Exploration Index is obtained.

Definition 3.9 [14] The Spatial Exploration Index of a player is given by:

$$SEI = \frac{\sum_i^N \sqrt{(x_i - x_m)^2 + (y_i - y_m)^2}}{N}, \tag{3.12}$$

where N represents the number of time instants for which the Spatial Exploration Index is being calculated, (x_m, y_m) the mean position of the player over the time period and (x_i, y_i) the position of the player on instant i.

3.5.2 Real Life Examples

The results obtained by three players in the SSG are presented in Table 3.17 with intervals of 30 s and represented in Fig. 3.8, with the results for the entire 3 min in Table 3.18.

Table 3.17 Values obtained for the Spatial Exploration Index in an SSG, for periods of 30 s

Average point (s)	Period of time (s)	Player 1 (m)	Player 2 (m)	Player 3 (m)
15	[0; 30[	40.5111	46.7643	43.3826
45	[30; 60[	35.2779	60.0238	43.8389
75	[60; 90[	44.5418	53.9460	49.5883
105	[90; 120[	41.1178	56.7801	48.8824
135	[120; 150[	38.7436	52.2909	47.4142
165	[150; 180]	44.9751	50.7073	47.0478

Table 3.18 Values obtained for the Spatial Exploration Index in an SSG, in the entire period of time of 3 min

Period of time (s)	Player 1 (m)	Player 2 (m)	Player 3 (m)
[0; 180]	40.8514	53.4120	46.6829

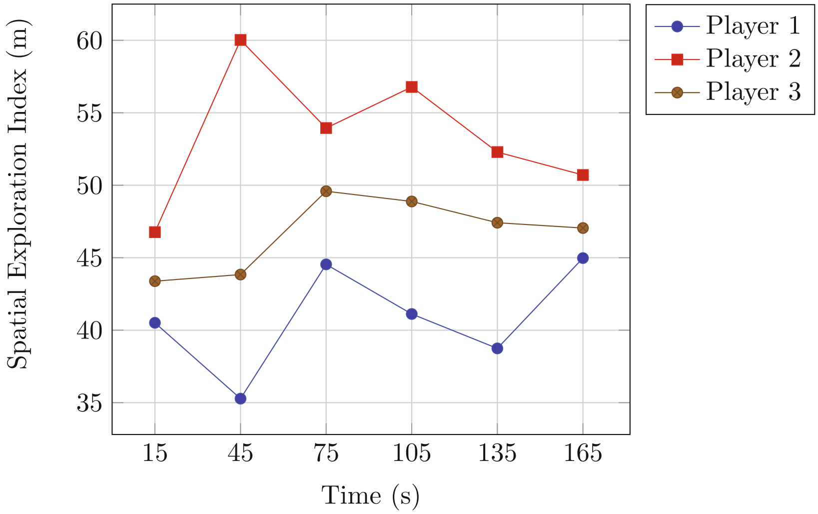

Fig. 3.8 Plotting of the values from Table 3.17, representing the Spatial Exploration Index in an SSG

These values can be compared to those obtained by another three players in the match, presented in Table 3.19 with intervals of 30 s and represented in Fig. 3.9, with the results for the entire 3 min in Table 3.20.

Table 3.19 Values obtained for the Spatial Exploration Index in a match, for periods of 30 s

Average point (s)	Period of time (s)	Player 1 (m)	Player 2 (m)	Player 3 (m)
15	[0; 30[	86.1651	88.7865	74.5967
45	[30; 60[	85.5437	82.2925	70.9353
75	[60; 90[	85.6321	84.0861	75.3764
105	[90; 120[	98.3602	98.5482	89.4149
135	[120; 150[	84.4817	82.2164	65.9566
165	[150; 180]	96.9700	96.0087	82.5427

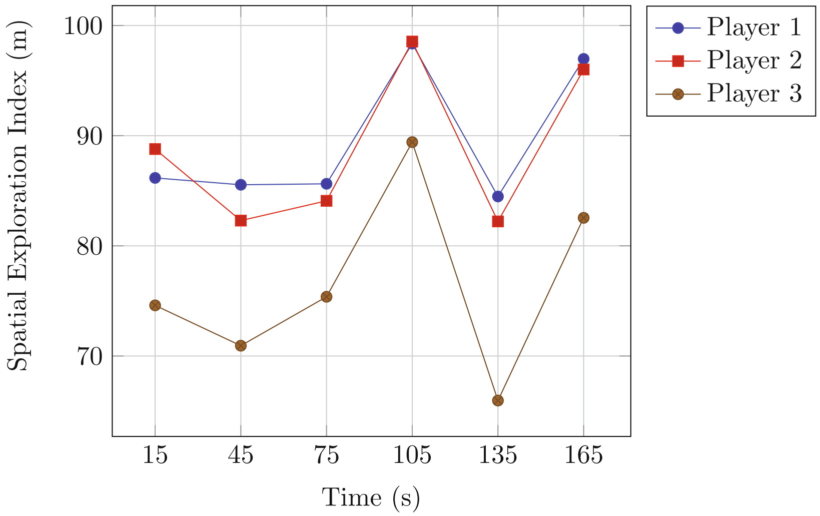

Fig. 3.9 Plotting of the values from Table 3.19, representing the Spatial Exploration Index in a match

Table 3.20 Values obtained for the Spatial Exploration Index in a match, in the entire period of time of 3 min

Period of time (s)	Player 1	Player 2	Player 3
[0; 180]	89.5296	88.6573	76.4681

3.5.3 General Interpretation

The Spatial Exploration Index was introduced as a novel measure to classify the exploration of a player's trajectory on the soccer pitch [14]. This measure uses the mean pitch position and the distance of all position time-series to identify how far a player goes beyond their "mean" point.

The values will vary based on the area of the soccer pitch and the positioning role of the player. However, this measure can be used to identify which formats of play and specific tasks may be used to improve the exploration of the soccer pitch or to restrict the movement of players. The classification of Spatial Exploration Index in different playing scenarios may help coaches make decisions about the occupied zone of the player and of ways to improve the movements on the soccer pitch.

References

1. Abt G, Lovell R (2009) The use of individualized speed and intensity thresholds for determining the distance run at high-intensity in professional soccer. J Sports Sci 27(9):893–898 PMID: 19629838
2. Akenhead R, Harley JA, Tweddle SP (2016) Examining the external training load of an English Premier League football team with special reference to acceleration. J Strength Cond Res 30(9)
3. Beezer RA (2008) A first course in linear algebra. Beezer, Tacoma
4. Bourdon PC, Cardinale M, Murray A, Gastin P, Kellmann M, Varley MC, Gabbett TJ, Coutts AJ, Burgess DJ, Gregson W, Cable NT (2017) Monitoring athlete training loads: consensus statement. Int J Sports Physiol Perform, 12(Suppl 2):S2–161–S2–170. PMID: 28463642
5. Carling C (2013) Interpreting physical performance in professional soccer match-play: should we be more pragmatic in our approach? Sports Med 43(8):655–663
6. Clemente FMC, Silva F, Martins F, Kalamaras D, Mendes R (2016) Performance analysis tool (pato) for network analysis on team sports: a case study of fifa soccer world cup 2014. J Sports Eng Technol 230(3):158–170
7. Couceiro MS, Clemente FM, Martins FML, Machado JAT (2014) Dynamical stability and predictability of football players: the study of one match. Entropy 16(2):645–674
8. Cummins C, Orr R, O'Connor H, West C (2013) Global positioning systems (gps) and microtechnology sensors in team sports: a systematic review. Sports Med 43(10):1025–1042
9. da Costa IT, Garganta J, Greco PJ, Mesquita I, Seabra A (2010) Influence of relative age effects and quality of tactical behaviour in the performance of youth soccer players. Int J Perform Anal Sport 10:82–97
10. Duarte R, Araújo D, Correia V, Davids K (2012) Sports teams as superorganisms: implications of sociobiological models of behaviour for research and practice in team sports performance analysis. Sport. Med. 42(8):633–642
11. Duarte R, Araújo D, Folgado H, Esteves P, Marques P, Davids K (2013) Capturing complex, non-linear team behaviours during competitive football performance. J Syst Sci Complex 26(1):62–72
12. Esteves PT, Araújo D, Davids K, Vilar L, Travassos B, Esteves C (2012) Interpersonal dynamics and relative positioning to scoring target of performers in 1 versus 1 sub-phases of team sports. J Sports Sci 30(12):1285–1293 PMID: 22852826
13. 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)
14. Gonçalves B, Esteves P, Folgado H, Ric A, Torrents C, Sampaio J (2017) Effects of pitch area-restrictions on tactical behavior, physical and physiological performances in soccer large-sided games. J. Strength Cond. Res. ahead-of-p
15. Malone JJ, Lovell R, Varley MC, Coutts AJ (2017) Unpacking the black box: applications and considerations for using GPS devices in sport. Int J Sports Physiol Perform 12(Suppl 2):S2–18–S2–26. PMID: 27736244
16. Pincus SM, Gladstone IM, Ehrenkranz RA (1991) A regularity statistic for medical data analysis. J Clin Monit 7(4):335–345
17. Salmon G (1865) A treatise on the analytic geometry of three dimensions. Hodges, Smith, and Company, Dublin
18. Silva P, Aguiar P, Duarte R, Davids K, Araújo D, Garganta J (2014) Effects of pitch size and skill level on tactical behaviours of Association Football players during small-sided and conditioned games. Int J Sports Sci Coach 9(5):993–1006. Online date: Monday, December 22, 2014
19. Villate JE (2013) Dinâmica e sistemas dinâmicos

Chapter 4
Metrics to Measure the Center of the Team

Abstract This chapter discusses a set of metrics involving the Geometrical Center (or centroid) of one or both teams. Each section describes a different metric, including the associated formulae and definitions, representations if valid, and an interpretation on what the metric can convey the user. The following measures will be presented: Geometrical Center; Longitudinal and Lateral Inter-team Distances; Time Delay between teams' movements and Coupling Strength. The case studies presented involve two five-player teams in an SSG considering only the space of half pitch (68 m goal-to-goal and 52 m side-to-side) and another eleven-player team in a match considering the space of the entire field (106.744 m goal-to-goal and 66.611 m side-to-side) even though only playing in half pitch.

Keywords Position data · Georeferencing · uPATO · Soccer · Collective behavior · Centroid

4.1 Geometrical Center

4.1.1 Basic Concepts

The average of the positions of all players of a team results in the Geometrical Center of the team, as the position that is center to the polygon formed by all of the team's players.

Definition 4.1 [9] The Geometrical Center of a team, $C(i)$, for any given instant i, is given by the following equation:

$$C(i) = \left(\frac{\sum_k^N p_{xk}(i)}{N}, \frac{\sum_k^N p_{yk}(i)}{N} \right) \tag{4.1}$$

where N represents the number of players in the team, $p_{xk}(i)$ the position along the longitudinal axis for player k in instant i, and $p_{yk}(i)$ the position along the vertical axis for player k in instant i.

F.M. Clemente et al., *Computational Metrics for Soccer Analysis*, SpringerBriefs in Applied Sciences and Technology, https://doi.org/10.1007/978-3-319-59029-5_4

4.1.2 Real Life Examples

The results obtained by both teams in the SSG and by another team in the match are presented in Table 4.1 with intervals of 30 s and for the entire 3 min in Table 4.2.

A screenshot of a representation of the Geometrical Center captured from the uPATO software is displayed in Fig. 4.1.

Table 4.1 Values obtained for the Geometrical Center of both teams in an SSG and one team in a match, for periods of 30 s

Period of time (s)	SSG		Match
	Team A (m)	Team B (m)	Team A (m)
[0; 30[	(33.1874, 27.8405)	(27.8414, 20.2657)	(78.3172, 36.8925)
[30; 60[	(30.5711, 32.4622)	(22.5556, 25.7821)	(69.1305, 36.7852)
[60; 90[	(36.2524, 28.3947)	(33.6091, 22.0747)	(73.8001, 37.4438)
[90; 120[	(34.4705, 34.2036)	(29.9851, 29.4189)	(92.7840, 32.9777)
[120; 150[	(33.3639, 28.6499)	(28.7977, 20.0259)	(71.7747, 36.8492)
[150; 180]	(32.6036, 34.1118)	(27.4019, 27.4541)	(86.3340, 36.1387)

Table 4.2 Values obtained for the Geometrical Center of both teams in an SSG and one team in a match, in the entire period of time of 3 min

Period of time (s)	SSG		Match
	Team A (m)	Team B (m)	Team A (m)
[0; 180]	(33.4010, 30.9450)	(28.3535, 24.1718)	(78.6921, 36.1801)

Fig. 4.1 Screenshot of the uPATO game animation with the representation and values of the Geometrical Center visible for both teams in the example SSG

4.1.3 General Interpretation

The Geometrical Center represents the middle point of a team [3]. Centroid, wcentroid or team's center have been also used as synonymous of Geometrical Center [5, 7, 9]. The Geometrical Center based on the Euclidian distance of the dots (players) provides useful information about the oscillation of the middle point of the team during the match and in specific circumstances [1]. In some cases the Geometrical Center has been used to identify the intra- and inter-team coordination tendencies in a temporal series [6], namely to monitor the in- and anti-phase relationships between Geometrical Centers from both teams [1, 3, 6]. In such analysis, some studies suggested that non-synchronization of geometrical centers is associated with critical moments [1, 5].

The Geometrical Center can be used by coaches to identify the specific position of the middle point of the team during defensive and attacking moments and associate such point with the Geometrical Center of the opponent team. Moreover, it can be used to control the distances of the farthest players to the center of the team. The association of the Geometrical Center with the oscillation of the ball can be also useful to check the capacity of the team to move collectively based on the position of the ball.

4.2 Longitudinal and Lateral Inter-team Distances

4.2.1 Basic Concepts

The different inter-team distances can be calculated through the difference between the positions of the geometrical centers of the two teams.

Definition 4.2 [6] The instantaneous Longitudinal Inter-team Distance can be calculated by the following equation:

$$iLoID = |x_{t1} - x_{t2}|, \tag{4.2}$$

where $x_{tk}, k = 1, 2$, represents the x coordinate of the Geometrical Center of team k on the calculated instant.

Definition 4.3 [6] The instantaneous Lateral Inter-team Distance is given by the following equation:

$$iLaID = |y_{t1} - y_{t2}|, \tag{4.3}$$

where $y_{tk}, k = 1, 2$, represents the y coordinate of the Geometrical Center of team k on the calculated instant.

Definition 4.4 [6] The instantaneous total Inter-team Distance is given by the following equation:

$$iID = \sqrt{(x_{t1} - x_{t2})^2 + (y_{t1} - y_{t2})^2}, \tag{4.4}$$

where (x_{tk}, y_{tk}), $k = 1, 2$, represents the coordinates of the Geometrical Center of team k on the calculated instant.

Definition 4.5 [6] The average Inter-team Distance can be calculated through the following equation:

$$\overline{ID} = \frac{\sum_{t=1}^{N} iID(t)}{N}, \tag{4.5}$$

where $iID(t)$ is the Inter-team Distance on a given instant t, and N the total of measured time instants. The same equation can be adapted for the average Longitudinal and Lateral Inter-team Distances, by replacing $iID(t)$ with $iLoID(t)$ or $iLaID(t)$, respectively.

4.2.2 Real Life Examples

The results obtained by a player in the SSG are presented in Table 4.3 with intervals of 30 s and for the entire 3 min in Table 4.4.

A screenshot of a representation of the total Inter-team Distance captured from the uPATO software is displayed in Fig. 4.2.

No results are presented for the match because this metric required the existence of data from both teams, which is not available in the evaluated match.

Table 4.3 Values obtained for the Longitudinal and Lateral Inter-team Distances in an SSG, for periods of 30 s

Period of time (s)	x axis (m)	y axis (m)	Both axes (m)
[0; 30[	5.3484	7.5770	9.2713
[30; 60[	8.0405	6.6773	10.4342
[60; 90[	3.4501	6.3190	6.8505
[90; 120[	4.4987	4.7876	6.5583
[120; 150[	4.5607	8.6208	9.7583
[150; 180]	5.2162	6.6662	8.4489

Table 4.4 Values obtained for the Longitudinal and Lateral Inter-team Distances in an SSG, in the entire period of time of 3 min

Period of time (s)	x axis (m)	y axis (m)	Both axes (m)
[0; 180]	5.1857	6.7746	8.4471

Fig. 4.2 Screenshot of the uPATO showing an example game animation with the representation and values of the Inter-team Distance

4.2.3 General Interpretation

The association between geometrical centers of both teams can be used to classify the synchronization during positional attacking, counter attack or defensive pressure. Longitudinal and Lateral Inter-team Distances represent the distance between Geometrical Centers of both teams in the two axes [6]. In the study conducted by the authors of this measure [6] a 3-sec window was implemented to monitor the capacity of this measure to anticipate critical match events.

The variability of the distance between both Geometrical Centers depends on contextual factors. However, this can be useful to assess which game situations approach or distance teams. Moreover, coaches can use such measure to classify the influence of specific small-sided and conditioned games in the spatio-temporal relationship. The classification of specific defensive (defensive 'block' or defensive transition) or attacking moments (positional attack or counter-attack) can be also made using this measure of distance in longitudinal and lateral axes.

4.3 Time Delay Between Teams' Movements

4.3.1 Basic Concepts

The Time Delay between two teams is the time difference a team takes to adjust to a change of positioning of the other team. This is approached by measuring the time it takes for the Geometrical Center of a team to approach the other team's center after its movement.

Definition 4.6 [2, 8] Given a time-series with N measurements, where gm_{t1} and gm_{t2} sets of points representing the Geometrical Center of each team across the time-series, a window size w, a time lag τ on the integer interval $-\tau_{max} \leq \tau \leq \tau_{max}$ and a time index i, a pair of time windows, Wx and Wy can be selected from the time-series, for each coordinate of the Geometrical Center. The time windows are given by:

$$Wx = \begin{cases} \{x_i, x_{i+1}, x_{i+2}, \ldots, x_{i+w_{max}}\} & \text{if } \tau \leq 0 \\ \{x_{i-\tau}, x_{i+1-\tau}, x_{i+2-\tau}, \ldots, x_{i+w_{max}-\tau}\} & \text{if } \tau > 0 \end{cases} \tag{4.6}$$

$$Wy = \begin{cases} \{y_{i+\tau}, y_{i+1+\tau}, y_{i+2+\tau}, \ldots, y_{i+w_{max}+\tau}\} & \text{if } \tau < 0 \\ \{y_i, y_{i+1}, y_{i+2}, \ldots, y_{i+w_{max}}\} & \text{if } \tau > 0 \end{cases} \tag{4.7}$$

Definition 4.7 [2, 8] The cross-correlation between the time windows can be defined as:

$$r(Wx, Wy) = \frac{1}{w_{max}} \sum_{i=1}^{w_{max}} \frac{(Wx_i - \mu(Wx))(Wy_i - \mu(Wy))}{sd(Wx)sd(Wy)}, \tag{4.8}$$

where $\mu(Wx)$ and $\mu(Wy)$ are the mean of Wx and Wy, and $sd(Wx)$ and $sd(Wy)$ are the standard deviation of Wx and Wy.

Definition 4.8 [8] Given the cross-correlation values for the different time windows of τ lag, the Time Delay for each instant is given by the following equation:

$$TD(i) = max\big(r(Wx, Wy)(i)\big), \tag{4.9}$$

where i represents the time instant of the time-series of N length.

To better represent the tendency of the time delays over a period of time, the graphs exemplified in Fig. 4.3 are created through the use of Fourier Series and linear regressions.

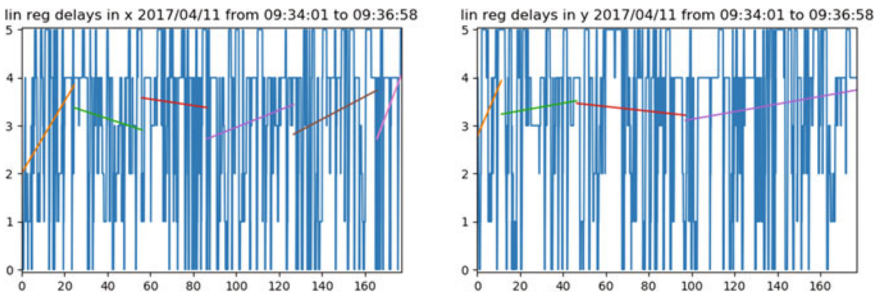

Fig. 4.3 Example of the graph illustrating the time delays between two teams and their tendency during a period of a game. The values presented in both axes are in seconds

Definition 4.9 [4] Given a set of time delays of length N, the graphical representation for the trend estimate of the time delay between the teams is calculated through the following steps:

(i) Define the Fourier Series with n terms:

$$q_n(t) = a_0 + \sum_{j=1}^{n} \left(a_j cos\left(\frac{j2\pi}{T} t\right) + b_j sin\left(\frac{j2\pi}{T} t\right) \right); \tag{4.10}$$

(ii) Determine the real roots of the first derivative:

$$t_1, t_2, \ldots, t_k \text{ such that } q'(t_1) = 0, \ldots, q'(t_k) = 0; \tag{4.11}$$

(iii) The segmented trend estimate is set to:

$$T(t) = \alpha_j + \beta_j(t), \tag{4.12}$$

for $t_{j-1} < t < t_j$, $t_0 = 0$ and where $j = 1, \ldots, k+1$.

4.3.2 Real Life Examples

The results obtained in the SSG are presented in Table 4.5 with intervals of 30 s and for the entire 3 min in Table 4.6.

Table 4.5 Values obtained for the Time Delay between teams' movements in an SSG, for periods of 30 s

Period of time (s)	x axis (s)	y axis (s)
[0; 30[	3	3
[30; 60[	2	3
[60; 90[	3	3
[90; 120[	2	3
[120; 150[	3	3
[150; 180]	3	3

Table 4.6 Values obtained for the Time Delay between teams' movements in an SSG, in the entire period of time of 3 min

Period of time (s)	x axis (s)	y axis (s)
[0; 180]	3	3

No results are presented for the match because this metric required the existence of data from both teams, which is not available in the evaluated match.

4.3.3 General Interpretation

The Time Delay can be measured in longitudinal (goal-to-goal) and lateral (side-to-side) directions. This measure assesses the existing Time Delay between both teams' movements using the Geometrical Center [8]. The measure quantifies the delay of teams in adjusting to each other's movements in both axes [8]. In the study that proposed this measure it was found that in the majority ($\approx$80%) of the cases the teams adjusted the geometrical centers in less than 0.5 s in the longitudinal axis [8]. More time was verified in lateral axis [8].

This measure can be used to identify longer periods (>0.5 s) of no-synchronization that can result in critical moments (e.g., shots, goals, counter-attacks). The association of shots/goals suffered and made can be performed to improve the usefulness of this measure to anticipate critical events. Moreover, this measure can be used to identify the evolution of synchronism in different periods of the match. Finally, coaches may use this information to design games that promote smaller or longer periods of delay and to adjust the team with real playing scenarios.

4.4 Coupling Strength

4.4.1 Basic Concepts

The Coupling Strength is calculated as the cross-correlation between the movement of the two teams, according to their Geometrical Centers.

Definition 4.10 [2, 8] Given a time-series with N measurements, where gm_{t1} and gm_{t2} sets of points representing the Geometrical Center of each team across the time-series, a window size w_{max}, a time lag τ on the integer interval $-\tau_{max} \leq \tau \leq \tau_{max}$ and a time index i, a pair of time windows, Wx and Wy can be selected from the time-series, for each coordinate of the Geometrical Center. The time windows are given by:

$$Wx = \begin{cases} \{x_i, x_{i+1}, x_{i+2}, \ldots x_{i+w_{max}}\} & \text{if } \tau \leq 0 \\ \{x_{i-\tau}, x_{i+1-\tau}, x_{i+2-\tau}, \ldots x_{i+w_{max}-\tau}\} & \text{if } \tau > 0 \end{cases} \tag{4.13}$$

$$Wy = \begin{cases} \{y_{i+\tau}, y_{i+1+\tau}, y_{i+2+\tau}, \ldots y_{i+w_{max}+\tau}\} & \text{if } \tau < 0 \\ \{y_i, y_{i+1}, y_{i+2}, \ldots y_{i+w_{max}}\} & \text{if } \tau > 0 \end{cases} \tag{4.14}$$

Definition 4.11 [2, 8] From the time windows, the cross-correlation between the time windows can be defined as:

$$r(Wx, Wy) = \frac{1}{w_{max}} \sum_{i=1}^{w_{max}} \frac{(Wx_i - \mu(Wx))(Wy_i - \mu(Wy))}{sd(Wx)sd(Wy)}, \tag{4.15}$$

where $\mu(Wx)$ and $\mu(Wy)$ are the mean of Wx and Wy, and $sd(Wx)$ and $sd(Wy)$ are the standard deviation of Wx and Wy. The Coupling Strength between the teams is the cross-correlation value at zero-lags, when $\tau = 0$.

4.4.2 Real Life Examples

The results obtained in the SSG are presented in Table 4.7 with intervals of 30 s and for the entire 3 min in Table 4.8.

No results are presented for the match because this metric required the existence of data from both teams, which is not available in the evaluated match.

4.4.3 General Interpretation

The Coupling Strength was proposed to quantify the degree of coordination/ synchronization of both teams' movements in both axes (longitudinal and lateral) [8]. This measure represents an association between the Geometrical Centers of both teams. In the unique study conducted with this measure, 80% of the time it was verified a

Table 4.7 Values obtained for the Coupling Strength in an SSG, for periods of 30 s

Period of time (s)	x axis (s)	y axis (s)
[0; 30[	0.8573	0.5026
[30; 60[	0.9769	0.8284
[60; 90[	0.8948	0.8316
[90; 120[	0.8810	0.6271
[120; 150[	0.1766	0.5366
[150; 180]	0.2692	0.5425

Table 4.8 Values obtained for the Coupling Strength in an SSG, in the entire period of time of 3 min

Period of time (s)	x axis (s)	y axis (s)
[0; 180]	0.6673	0.6419

Coupling Strength between 0.9 and 1 s, in longitudinal axis [8]. Smaller percentage of the time was verified in the lateral axis. Therefore, the Coupling Strength showed that both teams were highly synchronous for most of the match time [8].

This measure can be used to identify the time spent in synchronization between teams and to characterize the capacity of a team to adjust to the other. Moreover, can be used to identify critical moments of the match that may contribute to increase or decrease the Coupling Strength of the teams. Finally, different small-sided and conditioned games lead to different Coupling Strengths and for that reason coaches can use such information to work the capacity of the team to synchronize with other or only to test the development of Coupling Strength of two teams during the match.

References

1. Bartlett R, Button C, Robins M, Dutt-Mazumder A, Kennedy G (2012) Analysing team coordination patterns from player movement trajectories in football: methodological considerations. Int J Perform Anal Sport 12(2):398–424
2. Boker SM, Rotondo JL, Xu M, King K (2002) Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series. Psychol Methods 7(3):338
3. Clemente FM, Couceiro MS, Martins FML, Mendes RS, Figueiredo AJ (2014) Intelligent systems for analyzing soccer games: the weighted centroid. Ing. e Investig 34(3):70–75
4. Vicente MF, Martins F, Mendes R, Dias G, Fonseca J (2010) A method for segmented-trend estimate and geometric error analysis in motor learning. In: Proceedings of the First International conference on mathematical methods (MME10), pp 433–442 .
5. Frencken W, Lemmink K, Delleman N, Visscher C (2011) Oscillations of centroid position and surface area of football teams in small-sided games. Eur J Sport Sci 11(4):215–223
6. Frencken W, de Poel H, Visscher C, Lemmink K (2012) Variability of inter-team distances associated with match events in elite-standard soccer. J Sports Sci 30(12):1207–1213 PMID:22788797
7. Lames M, Ertmer J, Walter F (2010) Oscillations in football-order and disorder in spatial interactions between the two teams. Int J Sport Psychol 41(4):85
8. Silva P, Vilar L, Davids K, Araújo D, Garganta J (2016) Sports teams as complex adaptive systems: manipulating player numbers shapes behaviours during football small-sided games. SpringerPlus 5(1):191
9. Yue Z, Broich H, Seifriz F, Mester J (2008) Mathematical analysis of a football game. part i: individual and collective behaviors. Stud Appl Math 121(3):223–243

Chapter 5
Measuring the Dispersion of the Players

Abstract The purpose of this chapter is to introduce the concepts of dispersion in the aim of soccer analysis. A set of different measures have been proposed to identify the level of dispersion between teammates and between opponents. Based on that, a summary of the dispersion measures, definitions, interpretation and graphical visualization will be presented on this chapter. The measures of Stretch Index, Surface Area, Team Length and Team Width and lpwratio will be introduced throughout the chapter. The case studies presented involve two five-player teams in an SSG considering only the space of half pitch (68 m goal-to-goal and 52 m side-to-side) and another eleven-player team in a match considering the space of the entire field (106.744 m goal-to-goal and 66.611 m side-to-side) even though only playing in half pitch.

Keywords Position data · Georeferencing · uPATO · Soccer · Collective behavior · Team's dispersion

5.1 Stretch Index

5.1.1 Basic Concepts

The Stretch Index of a team is calculated as the average distance of a team's players and the Geometrical Center, giving a notion of the compactness of the team.

Definition 5.1 [1] The Stretch Index, considering both axes, at a given instant t can be calculated by:

$$SI(t) = \frac{\sum_k^N \sqrt{(p_{xk}(t) - C_x(t))^2 + p_{yk}(t) - C_y(t))^2}}{N}, \tag{5.1}$$

where $C(t)$ represents the Geometrical Center of the team, N the number of players in the team, $P_{xk}(t)$ the position along the longitudinal axis for player k at instant t, and $P_{yk}(t)$ the position along the vertical axis for player k in instant t.

F.M. Clemente et al., *Computational Metrics for Soccer Analysis*, SpringerBriefs in Applied Sciences and Technology, https://doi.org/10.1007/978-3-319-59029-5_5

Definition 5.2 [1] The Stretch Index on a single axis, for a given instant t, is given by the following expression:

$$SI_x(t) = \sum_{k}^{N} |(p_{xk}(t) - C_x(t))^2)|, \tag{5.2}$$

where $C_x(t)$ represents the x coordinate for the Geometrical Center at instant t and $P_{xk}(t)$ the position along the longitudinal axis for player k at instant t.

Remark 5.1 The same formula is applicable for the calculation along the vertical axis, only replacing $P_{xk}(t)$ for $P_{yk}(t)$ and $C_x(t)$ for $C_y(t)$.

Definition 5.3 [1] The average values of the Stretch Index both for the coordinates, along the two axes and for each axis separately, is given by the following equation:

$$\overline{SI} = \frac{\sum_k^N SI(t)}{N_t}, \tag{5.3}$$

where N_t represents the total number of time instants measured.

Remark 5.2 This same formula is applicable for $\overline{SI_x}$ and $\overline{SI_y}$, by replacing $SI(t)$ for $SI_x(t)$ or $SI_y(t)$, respectively.

5.1.2 Real Life Examples

The results obtained by a player from Team A in the SSG are presented in Table 5.1 with intervals of 30 s and for the entire 3 min in Table 5.2.

The results obtained by a player from Team B in the SSG are presented in Table 5.3 with intervals of 30 s and for the entire 3 min in Table 5.4.

A screenshot of a representation of the Stretch Index captured from the uPATO software is displayed in Fig. 5.1.

Table 5.1 Values obtained for the Stretch Index of Team A in an SSG, for periods of 30 s

Period of time (s)	x axis (m)	y axis (m)	Both axes (m)
[0; 30[	22.8927	5.6728	24.3221
[30; 60[	43.8870	18.9101	49.7405
[60; 90[	35.6433	10.6641	39.1172
[90; 120[	30.2275	12.5269	34.8085
[120; 150[	36.0131	14.0653	41.0395
[150; 180]	20.1069	5.3325	21.6990

Table 5.2 Values obtained for the Stretch Index of Team A in an SSG, in the entire period of time of 3 min

Period of time (s)	x axis (m)	y axis (m)	Both axes (m)
[0; 180]	31.4557	11.1946	35.1151

Table 5.3 Values obtained for the Stretch Index of Team B in an SSG, for periods of 30 s

Period of time (s)	x axis (m)	y axis (m)	Both axes (m)
[0; 30[	23.5324	15.6531	29.5884
[30; 60[	16.7984	13.6148	24.0102
[60; 90[	22.6443	16.5140	29.8536
[90; 120[	14.9087	12.3292	21.4474
[120; 150[	23.3738	19.5423	32.5294
[150; 180]	20.0127	16.9829	27.7828

Table 5.4 Values obtained for the Stretch Index of Team B in an SSG, in the entire period of time of 3 min

Period of time (s)	x axis (m)	y axis (m)	Both axes (m)
[0; 180]	20.2070	15.7707	27.5298

Fig. 5.1 Screenshot of the uPATO showing an example game animation with the representation and values of the Stretch Index for both teams

5.1.3 General Interpretation

Stretch Index was introduced in basketball to measure the expansion and contraction of space, in both axes (longitudinal and lateral), demonstrated by a team during a match [2]. This measure represents the mean deviation of each teammate to the geometrical center [2, 3]. Stretch Index is also known as radius [4].

Dispersion of the players depends on contextual variables and mostly on the moment of the game (with or without possession of the ball) [5]. Dispersion is greater in attacking moments (with possession of the ball) and smaller in defensive pressure (without possession of the ball), in the case of soccer [6]. This follows the main idea that in attacking moments it is necessary to spread the players to attract the opponents towards the outside of the middle and in defensive moments it is necessary to keep the teammates closer to guarantee fewer spaces for opponent's penetration [7].

This measure works in longitudinal, lateral and/or global, thus different information can be used. In the case of the Stretch Index for the longitudinal axis it can be computed to measure specific situations of counter attacks in which a greater dispersion in goal-to-goal direction can be observed. In the other hand, greater dispersions are found in side-to-side during positional attack (ball circulation). Considering the defensive moments, both axes will drastically decrease in comparison to attacking moments. However, longitudinal dispersion can be used to classify the defensive pressure against positional attacks in which large values of dispersion may suggest that forward teammates are too far away from the defensive colleagues.

5.2 Surface Area

5.2.1 Basic Concepts

The Surface Area of a team is calculated as the area of the polygon defined as the convex polygon with the least number of vertices that can encompass all of the teams's players, and where the potential vertices are defined as the positions of the players.

Definition 5.4 [8, 9] Given a set of points, the following algorithm is applied to define the Convex Hull:

Algorithm 1: Convex Hull algorithm.

```
Create a simplex of d+1 points
for each facet F do
    for each unassigned point p do
        if p is above F then
            assign p to F's outside set;
        end
    end
end
for each facet F with a non-empty outside set do
    select the furthest point p of F's outside set
    initialize the visible set V to F
    for all unvisited neighbours N of facets in V do
        if p is above N then
            add N to V
        end
    end
    the set of horizon ridges H is the boundary of V
    for each ridge R in H do
        create a new facet from R and p
        link the new facet to its neighbours
    end
    for each new facet F' do
        for each unassigned point q in an outside set of a facet in V do
            if q is above F' then
                assign q to F''s outside set
            end
        end
    end
    delete the facets in V.
end
```

Definition 5.5 [9, 10] Given the coordinates of the n vertices that compose the Convex Hull of the team, the Surface Area is given by the following equation:

$$SA = \frac{|(x_1y_2 - y_1x_2) + (x_2y_3 - x_3y_2) + \cdots + (x_ny_1 - x_1y_n)|}{2}, \tag{5.4}$$

where (x_i, y_i) are the coordinates of the *ith* vertex of the Convex Hull.

5.2.2 *Real Life Examples*

The results obtained by both teams in the SSG and another team in the match are presented in Table 5.5 with intervals of 30 s and for the entire 3 min in Table 5.6.

A screenshot of a representation of the Surface Area captured from the uPATO software is displayed in Fig. 5.2.

Table 5.5 Values obtained for the Surface Area of both teams in an SSG and obtained from a team in a match, for periods of 30 s

Period of time (s)	Surface Area (m^2)		
	SSG		Match
	Team A	Team B	Team A
[0; 30[	83.0395	247.7091	485.1521
[30; 60[	550.1110	154.9699	282.3502
[60; 90[	263.9492	277.3812	352.4851
[90; 120[	320.9839	131.7475	135.3420
[120; 150[	397.8896	303.0914	420.1875
[150; 180]	72.6705	245.4415	228.5462

Table 5.6 Values obtained for the Surface Area of both teams in an SSG and obtained from a team in a match, in the entire period of time of 3 min

Period of time (s)	Surface Area (m^2)		
	SSG		Match
	Team A	Team B	Team A
[0; 180]	281.3740	226.6473	317.4336

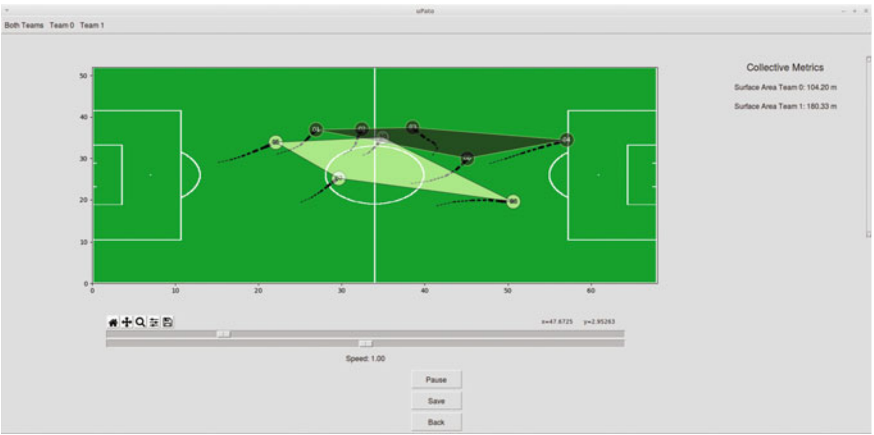

Fig. 5.2 Screenshot of the uPATO showing an example game animation with the representation and values of the surface area for both teams

5.2.3 General Interpretation

Surface Area represents the area of a polygon constituted by all teammates (dots) [11]. The measure can be defined as the total space covered by a team considering the area within the convex hull [9]. The Surface Area measures the contraction and expansion of teams across the soccer match, as does the Stretch Index [3]. However, this measure represents all the area covered by the team, while the Stretch Index only measures the mean deviation to the Geometrical Center of the team.

Typically, the area of the teams is significantly bigger in possession of the ball than without possession [3, 5]. Moreover, the variability of Surface Area decreases across the game [4, 12], thus suggesting a stabilization of the team in attacking and defensive moments.

In a study conducted in elite Spanish soccer teams values between 800 and 2800 m^2 in attacking and [1000; 2000] m^2 in defensive moments were found [13]. Values in attacking varied between 1638 and 1831 m^2 in possession of the ball and [1277; 1369] m^2 without possession in an elite Portuguese team [14].

The visualization of Surface Area can help coaches identify the space between sectors (defensive, middle and forward) and the dispersion of the team in specific moments. The generated triangulations can provide an immediate analysis of the collective behavior, particularly in moments of positional attack and defensive 'block'.

5.3 Team Length and Team Width

5.3.1 Basic Concepts

The Team Length and Width is defined by its most advanced and rear players for the length, and its rightmost and leftmost players for the length.

Definition 5.6 [4] Given a set of points for team player positioning along a time-series of length N, P, and where P_x and P_y represent the set of longitudinal and lateral coordinates for every player of the team on every measured time instant, the Team Length and Team Width on a given time instant i can be calculated as follows:

$$t_l(i) = max(P_x(i)) - min(P_x(i)) \tag{5.5}$$

$$t_w(i) = max(P_y(i)) - min(P_y(i)), \tag{5.6}$$

where t_l represents the Team Length, t_w represents the Team Width and $(P_x(i), P_y(i))$ represents the set of coordinates of the team's players in instant i.

5.3.2 Real Life Examples

The results obtained by both teams in the SSG are presented in Table 5.7 with intervals of 30 s and for the entire 3 min in Table 5.8.

The results obtained by the team in the match are presented in Table 5.9 with intervals of 30 s and for the entire 3 min in Table 5.10.

Table 5.7 Values obtained for the team width and team length of both teams in an SSG, for periods of 30 s

Period of time (s)	Team A		Team B	
	Team width (m)	Team length (m)	Team width (m)	Team length (m)
[0; 30[	25.2440	7.2677	30.7833	15.2246
[30; 60[	40.2148	18.7316	20.2804	15.5792
[60; 90[	36.7722	11.7441	28.5761	15.9253
[90; 120[	31.0945	15.2547	20.1710	14.5367
[120; 150[	37.0654	14.6651	27.0824	18.7516
[150; 180]	21.1475	6.6030	24.8196	19.4085

Table 5.8 Values obtained for the team width and team length of both teams in an SSG, in the entire period of time of 3 min

Period of time (s)	Team A		Team B	
	Team width (m)	Team length (m)	Team width (m)	Team length (m)
[0; 180]	31.9227	12.3758	25.2784	16.5720

Table 5.9 Values obtained for the team width and team length of a team in a match, for periods of 30 s

Period of time (s)	Team width (m)	Team length (m)
[0; 30[	23.2272	37.0941
[30; 60[	24.6797	21.2807
[60; 90[	26.5411	23.4404
[90; 120[	10.3152	28.7250
[120; 150[	26.3454	27.8867
[150; 180]	18.1004	23.2040

Table 5.10 Values obtained for the team width and team length of a team in a match, in the entire period of time of 3 min

Period of time (s)	Team width (m)	Team length (m)
[0; 180]	21.5368	26.9402

5.3.3 General Interpretation

The Team Length represents the maximum length of a team considering the minimum and maximum position of a player in the longitudinal (goal-to-goal) direction [4]. The same application is applied in the case of Team Width (side-to-side) [4]. Therefore, the length and width measures provide information about how stretched are the two farthest players of a team in longitudinal and lateral directions.

Coaches can use this information to understand the optimal distances to readjust some tactical tasks in training sessions based on usual length and width found in official matches. An interesting approach used the width and length to suggest specific sizes to work positional attack and counter-attack in small-sided games [15]. Moreover, coaches can compare the length and width in different moments of the match and identify in which moments the extreme size can be associated with critical moments (e.g., shots, goals).

5.4 Length per Width Ratio

5.4.1 Basic Concepts

The ratio between a team's length and width is the Length per Width Ratio (lpwratio) of that team.

Definition 5.7 [16] Given t_l and t_w as the Team Length and Team Width of a team on a given instant, the lpwratio of a team is given by the following equation:

$$lpwratio = \frac{t_l}{t_w}. \tag{5.7}$$

5.4.2 Real Life Examples

The results obtained by both teams in the SSG and the other team in the match are presented in Table 5.11 with intervals of 30 s and for the entire 3 min in Table 5.12.

A screenshot of a representation of the lpwratio captured from the uPATO software is displayed in Fig. 5.3.

5.4.3 General Interpretation

The lpwratio quantifies the relationship between the length (maximum distance between the two farthest players in longitudinal direction) and width (maximum

Table 5.11 Values obtained for the lpwratio of both teams in an SSG and a team in a match, for periods of 30 s

Period of time (s)	lpwratio		
	SSG		Match
	Team A	Team B	Team A
[0; 30[	0.2970	0.4971	1.6110
[30; 60[	0.4570	0.8310	0.8886
[60; 90[	0.3188	0.5621	0.8915
[90; 120[	0.4507	0.9115	2.9751
[120; 150[	0.3614	0.6977	1.1507
[150; 180]	0.3218	0.8415	1.3280

Table 5.12 Values obtained for the lpwratio of both teams in an SSG and a team in a match, in the entire period of time of 3 min

	lpwratio		
	SSG		Match
[0; 180]	0.3677	0.7238	1.4741

Fig. 5.3 Screenshot of the uPATO showing an example game animation with the representation and values of the width, length and lpwratio for both teams

distance between the two farthest players in lateral direction) during the match [16]. The authors of this measure argue that small variations of this measure may suggest a team's higher adherence to width and concentration on "principles of play" [16]. In the other hand, large variations of this ratio may suggest a more individual and less collectively coordinated approach to the soccer game [16].

Coaches can use this measure to classify the teams in different moments of the match. Some teams will tend to play in counter-attack, thus increasing the length and reducing the width. In the other hand, teams that opt to attack with circulation of the ball will have increased width and reduced the length. This relationship will help to understand some patterns of play. The same case can be applied for defensive moments in which teams that opt to defend in 'block' closer to the goal will present decreased length and teams that opt to extend the defensive 'block' for middle or forward zones of the pitch will have increased length.

References

1. Silva P, Vilar L, Davids K, Araújo D, Garganta J (2016) Sports teams as complex adaptive systems: manipulating player numbers shapes behaviours during football small-sided games. SpringerPlus 5(1):191
2. Bourbousson J, Sève C, McGarry T (2010) Space-time coordination dynamics in basketball: part 2. the interaction between the two teams. J Sport Sci 28(3):349–358
3. Bartlett R, Button C, Robins M, Dutt-Mazumder A, Kennedy G (2012) Analysing team coordination patterns from player movement trajectories in football: methodological considerations. Int J Perform Anal Sport 12(2):398–424
4. Duarte R, Araújo D, Folgado H, Esteves P, Marques P, Davids K (2013) Capturing complex, non-linear team behaviours during competitive football performance. J Sys Sci Complex 26(1):62–72
5. Clemente FM, Couceiro MS, Martins FML, Mendes RS, Figueiredo AJ (2013a) Measuring tactical behaviour using technological metrics: case study of a football game. Int J Sport Sci Coach 8(4):723–739
6. da Costa IT, Garganta J, Greco PJ, Mesquita I, Seabra A (2010) Influence of relative age effects and quality of tactical behaviour in the performance of youth soccer players. Int J Perform Anal Sport 10:82–97
7. Clemente FM, Martins F, Mendes R, Figueiredo A (2015) A systemic overview of football game: the principles behind the game. J Hum Sport Exerc 9(2):656–667
8. Barber CB, Dobkin DP, Huhdanpaa H (1996) The quickhull algorithm for convex hulls. ACM Trans Math Softw (TOMS) 22(4):469–483
9. Frencken W, Lemmink K, Delleman N, Visscher C (2011) Oscillations of centroid position and surface area of football teams in small-sided games. Eur J Sport Sci 11(4):215–223
10. Bourke P (1988) Calculating the area and centroid of a polygon. http://paulbourke.net/geometry/polygonmesh
11. Clemente FM, Couceiro MS, Martins FML, Mendes RS, Figueiredo AJ (2013b) Measuring collective behaviour in football teams: inspecting the impact of each half of the match on ball possession. Int J Perform Anal Sport 13(3):678–689
12. Moura FA, Martins LEB, Anido RO, Ruffino PRC, Barros RML, Cunha SA (2013) A spectral analysis of team mics and tactics in brazilian football. J Sport Sci 31(14):1568–1577 PMID: 23631771
13. 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 13:822–832
14. Clemente FM, Couceiro MS, Martins FML, Mendes RS, Figueiredo AJ (2014) Using collective metrics to inspect spatio-temporal relationships between football players. South African J Res Sport Phys Educ Recreat 36(2):47–59
15. Fradua L, Zubillaga A, Caro O, Iván Fernández-García A, Ruiz-Ruiz C, Tenga A (2013) Designing small-sided games for training tactical aspects in soccer: extrapolating pitch sizes from full-size professional matches. J Sports Sci 31(6):573–581
16. Folgado H, Lemmink KA, Frencken W, Sampaio J (2014) Length, width and centroid distance as measures of teams tactical performance in youth football. Eur J Sport Sci 14(sup1):S487–S492

Chapter 6
Measuring the Tactical Behavior

Abstract Tactical information can be determinant to use position data and measures in the aim of match analysis. By using information about collective behavior and tactics it is possible to re-organize tasks or even make decisions during matches. These measures are not limited to the space (as centroid or team's dispersion) but can also provide information on how teammates interact in the specificity of game and in line with tactical principles. Definitions, graphical visualization, interpretation and case-studies will be presented on this chapter for the following measures: Inter-player Context, Teams' Separateness, Directional Correlation Delay, Intra-team Coordination Tendencies, Sectorial Lines, Inter-axes of the team, Dominant Region, Major Ranges and Identification of Team's Formations. The case studies presented involve two five-player teams in an SSG considering only the space of half pitch (68 m goal-to-goal and 52 m side-to-side) and another eleven-player team in a match considering the space of the entire field (106.744 m goal-to-goal and 66.611 m side-to-side) even though only playing in half pitch.

Keywords Position data · Georeferencing · uPATO · Soccer · Collective behavior · Tactics

6.1 Inter-player Context

6.1.1 Basic Concepts

By taking into account the position of a player in regards to the goals and the adversary team positions, one can establish context between players in accordance to game situations.

Definition 6.1 [26] Given a time-series of length N of the positions of a player across the pitch, and the time-series representing the positions of the players of the opposing team, the following contexts are defined for each positional situation:

The player position is represented as p_{pos}, and the position of an opposing team player is represented by $p_{t2_{pos}}$.

F.M. Clemente et al., *Computational Metrics for Soccer Analysis*, SpringerBriefs
in Applied Sciences and Technology, https://doi.org/10.1007/978-3-319-59029-5_6

Algorithm 2: Inter-player context definition. "Player" is replaced by Rear, Intermediate or Advanced, depending on the role the player occupies.

```
if p_pos < min(p_t2_pos) then
    if Position of own goal is 0 then
        Player between advanced opponent and own goal
    end
    else
        if Position of own goal is the field width then
            Player between rear opponent and the opposing goal
        end
    end
end
if p_pos > max(p_t2_pos) then
    if Position of own goal is 0 then
        Player between rear opponent and the opposing goal
    end
    else
        if Position of own goal is the field width then
            Player between advanced opponent and own goal
        end
    end
end
else
    Player between advanced and rear opponent
end
```

6.1.2 General Interpretation

The Inter-player Context represents the positional changes of players with respect to all other players, thus being the relative position with respect to teammates and opponents [25]. The Inter-player Context proposed has nine possible contexts [25]: (i) "Rear teammate between advanced opponent and own goal"; (ii) "Intermediate teammate between advanced opponent and own goal"; (iii) "Advanced teammate between advanced opponent and own goal"; (iv) "Rear teammate between advanced and rear opponent"; (v) "Intermediate teammate between advanced and rear opponent"; (vi) "Advanced teammate between advanced and rear opponent"; (vii) "Rear teammate between rear opponent and the opposing goal"; (viii) "Intermediate teammate between rear opponent and the opposing goal"; (ix) "Advanced teammate between rare opponent and the opposing goal".

This measure can be used to classify the positioning of a given player considering the teammates and the opponents and to determine the percentage of time and frequency of time spent at different playing contexts. This measure can help coaches understand the influence of small-sided and constrained games or similar tasks in the development of specific tactical behavior of players. Moreover, the information about the player's contexts can provide opportunities to optimize the behavior in next occasions.

6.2 Teams' Separateness

6.2.1 Basic Concepts

Teams' Separateness is calculated as the sum of the distances between each of the team's players and their closest opponents.

Definition 6.2 [27] Given a time-series of length N with position data on each player of each team, and where each team is composed of N_{team} players, the Teams' Separateness, TS, can be calculated as follows:

$$TS = \sum_{i=1}^{N} min\Big(d(i)\Big), \tag{6.1}$$

such that $d(i) = \sqrt{(x_j(i) - x_k(i))^2 + (y_j(i) - y_k(i))^2}$, $j, k = 1, 2, 3, ..., N_{team}$, and where $(x_j(i), y_j(i))$ represents the positions of player j in instant i, where j is always a player of the first team, and $(x_k(i), y_k(i))$ represents the positions of player k in instant i, where k is always a player of the second team.

6.2.2 Real Life Examples

The results obtained by both teams in the SSG are presented in Table 6.1 with intervals of 30 s and for the entire 3 min in Table 6.2.

A screenshot of a representation of the Team's Separateness captured from the uPATO software is displayed in Fig. 6.1.

Table 6.1 Values obtained for the team's separateness of both teams in an SSG, for periods of 30 s

Period of time (s)	Team's separateness (m)	
	Team A	Team B
[0; 30[	7.5920	10.3459
[30; 60[	14.1655	9.4267
[60; 90[	9.0050	9.3900
[90; 120[	10.5678	8.9297
[120; 150[	11.1233	10.5824
[150; 180]	7.1063	9.8244

Table 6.2 Values obtained for the team's separateness of both teams in an SSG, in the entire period of time of 3 min

Period of time (s)	Team's separateness (m)	
	Team A	Team B
[0; 180]	10.4266	10.2518

Fig. 6.1 Screenshot of the uPATO game animation with the representation and values of the teams' separateness visible for both teams in the example SSG

No results are presented for the match because this metric required the existence of data from both teams, which is not available in the evaluated match.

6.2.3 General Interpretation

The Teams' Separateness quantifies the degree of free movement each team has available and estimates the amount of space separating the players of both teams [27]. This can be measured by using the average distance between all players and their closest opponent, being interpreted as the average radius of action free of opponents [27].

In the original study conducted in different SSG it were found no significant changes in the Teams' Separateness between different formats of the game (3 vs. 3, 4 vs. 4 and 5 vs. 5) [27]. Such results did not confirm the idea that the reduction of relative space of player decreases the distance values to immediate opponent players.

Teams' Separateness can be used by coaches to make decisions about which formats of play and pitch size must be designed to ensure greater or smaller free spaces without opponents. In specific circumstances such as in the last third of the pitch or in the scoring box it is required to develop the capacity to play in small spaces as best as possible. For these cases, the Teams' Separateness will be smaller and the coach may organize tasks that replicate such context. In the other hand, in the first third of the pitch there are more space to organize the circulation of the ball and for that reason the free space is greater and coaches may use such values to adjust training tasks according to the values of the Teams' Separateness.

6.3 Directional Correlation Delay

6.3.1 Basic Concepts

The Directional Correlation Delay is the delay that exists between the movements of a pair of players. It is calculated by selecting the delay where the correlation of players from the pair is maximum, within a selected time interval.

Definition 6.3 [21] The Directional Correlation function ($C_{ij}(\tau)$) in a given period of time, with a delay τ, between two players of a team, is calculated as follows:

$$C_{ij}(\tau) = \langle \overrightarrow{v_i}(t).\overrightarrow{v_j}(t+\tau)\rangle, \tag{6.2}$$

where $\overrightarrow{v_i}(t)$ is the normalized velocity vector of player i in instant t, and $\overrightarrow{v_j}(t+\tau)$ is the normalized velocity vector of player j in instant $t+\tau$. $\langle \ldots \rangle$ represents an average over time, for each instant t in that period of time.

Definition 6.4 The Directional Correlation delay (τ_{ij}^*) in a given period of time is determined by:

$$\tau_{ij}^* = \underset{\tau\in[-w,w]}{arg\ max}(C_{ij}(\tau)) \tag{6.3}$$

where $C_{ij}(\tau)$ is the Directional Correlation function and w is the number of seconds of the player j that should be analyzed before and after the current instant.

Remark 6.1 In terms of implementation, the time interval was converted from continuous to discrete:

$$\tau_{ij}^* = \underset{\tau\in\{-w,-w+\epsilon,\ldots,w-\epsilon,w\}}{arg\ max}(C_{ij}(\tau)) \tag{6.4}$$

using $w = 1$ and $\epsilon = 0.1$.

6.3.2 Real Life Examples

The results obtained by Team A in the SSG are presented for the entire 3 min in Table 6.3. Empty cells represent invalid possibilities (diagonal of the table) or repeated cases already presented in other cells (cell ij is equal to cell ji).

The results obtained by Team B in the SSG are presented for the entire 3 min in Table 6.4.

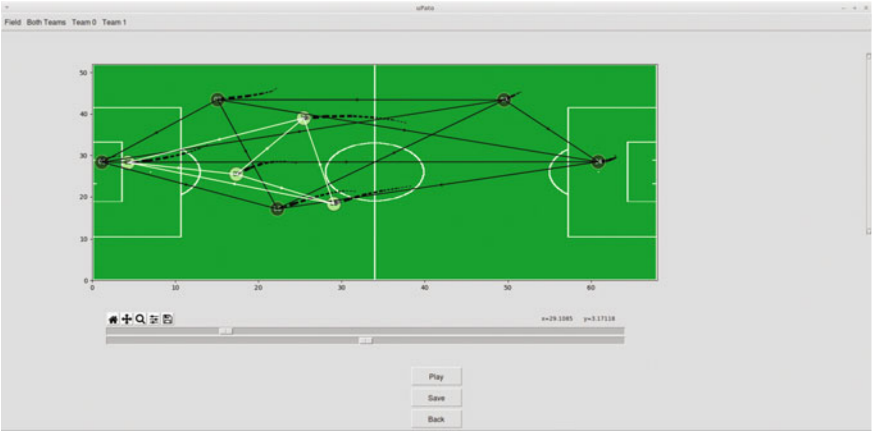

Fig. 6.2 Screenshot of the uPATO game animation with the representation and values of the directional correlation delay visible for both teams in the example SSG

These values can be compared to those obtained by another team in the match, presented for the entire 3 min in Table 6.5.

A screenshot of a representation of the Directional Correlation Delay captured from the uPATO software is displayed in Fig. 6.2.

Table 6.3 Values obtained for the directional correlation delay of Team A in an SSG, in the entire period of time of 3 min

	Player 1 (s)	Player 2 (s)	Player 3 (s)	Player 4 (s)	Player 5 (s)
Player 1		0.1	0.1	0.7	0.4
Player 2			0.7	0.4	−0.5
Player 3				−0.5	0.4
Player 4					−0.3
Player 5					

Table 6.4 Values obtained for the directional correlation delay of Team B in an SSG, in the entire period of time of 3 min

	Player 1 (s)	Player 2 (s)	Player 3 (s)	Player 4 (s)
Player 1		0.2	−0.1	−0.7
Player 2			−0.7	−0.2
Player 3				0.7
Player 4				

Table 6.5 Values obtained for the directional correlation delay of Team A in a match, in the entire period of time of 3 min

	Player 1 (s)	Player 2 (s)	Player 3 (s)	Player 4 (s)	Player 5 (s)	Player 6 (s)	Player 7 (s)	Player 8 (s)	Player 9 (s)
Player 1		0.8	−0.4	0.5	0.9	0.8	−0.1	0.2	0.3
Player 2			0.5	0.9	0.8	−0.1	0.2	0.3	−0.9
Player 3				0.8	−0.1	0.2	0.3	−0.9	−0.5
Player 4					0.2	0.3	−0.9	−0.5	−0.9
Player 5						−0.9	−0.5	0.9	−0.6
Player 6							0.9	−0.6	−0.8
Player 7								−0.8	0.5
Player 8									−0.9
Player 9									

6.3.3 General Interpretation

Directional Correlation Delay measures the delay between a movement and this "copied" movement from another player. This measure comes from the study of leaders in bird flocks [21]. The pairwise comparison allows to measure the directional leader-follower network during the match. The delay can either be positive or negative, with a positive delay meaning the first player is the leader, the one who performs the "original" movement, and a negative delay that it is the second player who is the leader.

In the same team it is possible to verify which players are more engaged with the teammates or with opponents and which ones act more independently of the remaining players. This can be interesting to classify the individual profile of players. Some of players may have the profile of leaders and others of followers.

This analysis can be determinant in youth teams to identify which teammates can promote the organization of the team. Moreover, in the case of the sectors (defensive, middle or forward) it is possible to identify the players that guide the colleagues in defensive and attacking movements.

6.4 Intra-team Coordination Tendencies

6.4.1 Basic Concepts

The Intra-team Coordination Tendencies of a team can be calculated by the percentage of time spent in-phase between each pair of players, which can be calculated through the relative phase between a pair of players on the same instant, which are then clustered through the k-means method into three clusters: high synchronization, intermediate synchronization and low synchronization.

Definition 6.5 [4] Given a time-series of length N, where each player has N position tuples of their position on each measured time instant, denoted as $x_i, i = 0, 1, \dots, N-1$, the Discrete Hilbert Transform (DHT) of this time-series, H_i, is a sequence $H_i, i = 0, 1, \dots, N-1$ defined by the following equation:

$$H_i = \begin{cases} \frac{2}{N}\sum_{\nu=0,2,4\dots} f_\nu coth(\nu - i)\frac{\pi}{N} & \text{if } i \text{ is odd} \\ \frac{2}{N}\sum_{\nu=1,3,5\dots} f_\nu coth(\nu - i)\frac{\pi}{N} & \text{if } i \text{ is even,} \end{cases} \tag{6.5}$$

where $coth(\alpha)$ is the hyperbolic cotangent.

Definition 6.6 [23] The relative phase between two signals, $\Phi(t)$, is given by the following formula:

$$\Phi(t) = arctan\left(\frac{H_1(t)s_2(t) - s_1(t)H_2(t)}{s_1(t)s_2(t) - H_1(t)H_2(t)}\right), \tag{6.6}$$

where $s_i(t)$ is the signal i and $H_i(t)$ its DHT.

Definition 6.7 [9] If the phase is between $-30°$ and $30°$, the signals are considered to be in-phase. The total time period in-phase, $\Delta\Phi$, between a player pair is given by the following equation:

$$\Delta\Phi = \sum_{i=1}^{t}[i \in \{0, 1, \ldots, t-1\} : -30° < \Phi(i) < 30°]. \tag{6.7}$$

Definition 6.8 The Intra-team Coordination Tendencies of a team is determined by:

$$ICT = \frac{\Delta\Phi}{\Delta t} * 100, \tag{6.8}$$

where Δt is the interval of time being analyzed.

Definition 6.9 [19] Given a set of player pairs and their in-phase time-periods, the k-means clustering algorithm classifies the Intra-Team Coordination Tendencies between pairs. The following formula specifies the objective of k-means clustering:

$$arg \quad min \sum_{i=1}^{k} \sum_{x \in S_i} ||x - \mu_i||^2, \tag{6.9}$$

where S_i are the sets formed by the clustering algorithm, and x the observations, in this case, the $\Delta\Phi$ of the player pairs.

Remark 6.2 In this case, considering we have three clusters: high, intermediate and low synchronization between players, $k = 3$.

6.4.2 Real Life Examples

The results obtained from Team A in the SSG are presented for the entire 3 min in Table 6.6.

The results obtained from Team B in the SSG are presented for the entire 3 min in Table 6.7.

These values can be compared to those obtained by another team in the match, presented for the entire 3 min in Table 6.8.

A screenshot of a representation of the Intra-team Coordination Tendencies in the x axis captured from the uPATO software is displayed in Fig. 6.3.

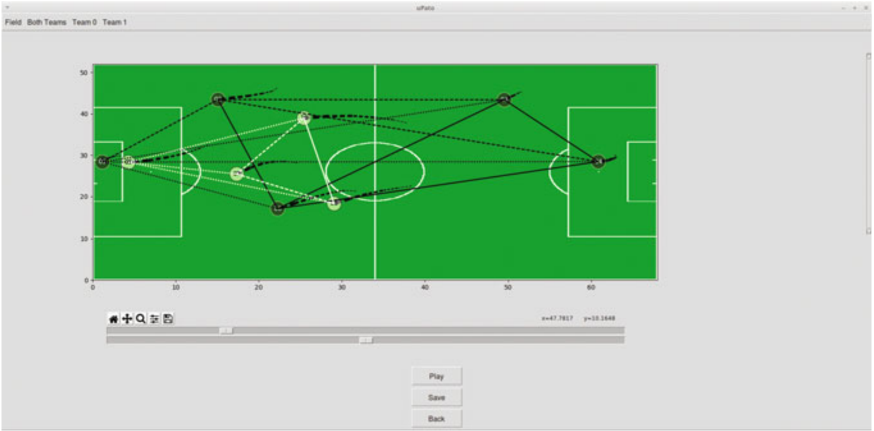

Fig. 6.3 Screenshot of the uPATO game animation with the representation and values of the intra-team coordination tendencies in the *x axis* visible for both teams in the example SSG

6.4.3 General Interpretation

The Intra-team Coordination Tendencies quantify the sharing of a common goal between pairs of players [9]. In this case, the percentage of time spent between

Table 6.6 Values of time spent in-phase between each pair of players ($\Delta\Phi$), used in the calculation of the intra-team coordination tendencies of Team A in the x axis in an SSG, in the entire period of time of 3 min

	Player 1	Player 2	Player 3	Player 4	Player 5
Player 1		153.4	130.9	137.3	138.1
Player 2			137.3	138.1	146.2
Player 3				146.2	143.7
Player 4					179.4
Player 5					

Table 6.7 Values of time spent in-phase between each pair of players ($\Delta\Phi$), used in the calculation of the intra-team coordination tendencies of Team B in the x axis in an SSG, in the entire period of time of 3 min

	Player 1	Player 2	Player 3	Player 4
Player 1		159.4	176.9	178.0
Player 2			178.0	159.5
Player 3				155.4
Player 4				

Table 6.8 Values of time spent in-phase between each pair of players ($\Delta\Phi$), used in the calculation of the intra-team coordination tendencies of Team A in the x axis in a match, in the entire period of time of 3 min

	Player 1	Player 2	Player 3	Player 4	Player 5	Player 6	Player 7	Player 8	Player 9
Player 1		180.1	180.1	180.1	180.1	180.1	180.1	180.1	180.1
Player 2			180.1	180.1	180.1	180.1	180.1	180.1	180.1
Player 3				180.1	180.1	180.1	180.1	180.1	180.1
Player 4					180.1	180.1	180.1	180.1	180.0
Player 5						180.1	180.1	180.0	180.1
Player 6							180.0	180.1	180.1
Player 7								180.1	180.1
Player 8									180.1
Player 9									

$-30°$ and $30°$ of relative phase was used to classify the sharing goals [9]. Intra-team Coordination Tendencies are measured at lateral and longitudinal directions.

The study of Intra-Team Coordination performed in 20 professional players revealed that central and lateral defenders were highly synchronized in lateral direction but little in longitudinal direction [9]. Specific clusters may emerge from context and the sectors of the team may contribute to high dependency between teammates. The coordination between sectors cannot be as high as intra-sector.

This measure can quantify the capacity of players to be coordinate in specific moments of the match and to classify the dependency between teammates. The intra-sector coordination in lateral or longitudinal displacement is highly important to build a solid team. The Intra-team Coordination Tendencies can be used to measure a capacity to move synchronously. Moreover, coaches can make decisions about which formation and group of players may contribute to a higher level of synchronization.

6.5 Sectorial Lines

6.5.1 Basic Concepts

The Sectorial Lines are lines that represent the regions of action of each role: defenders, midfielders and attackers. The line is the first degree polynomial that minimizes the Root Mean Squared Error (RMSE) between the polynomial and the players' positions [5]. A line is defined by the expression $y = \alpha + \beta x$, which can be calculated using a simple linear regression applied to the coordinates of the players performing that role, given in Definition 6.10.

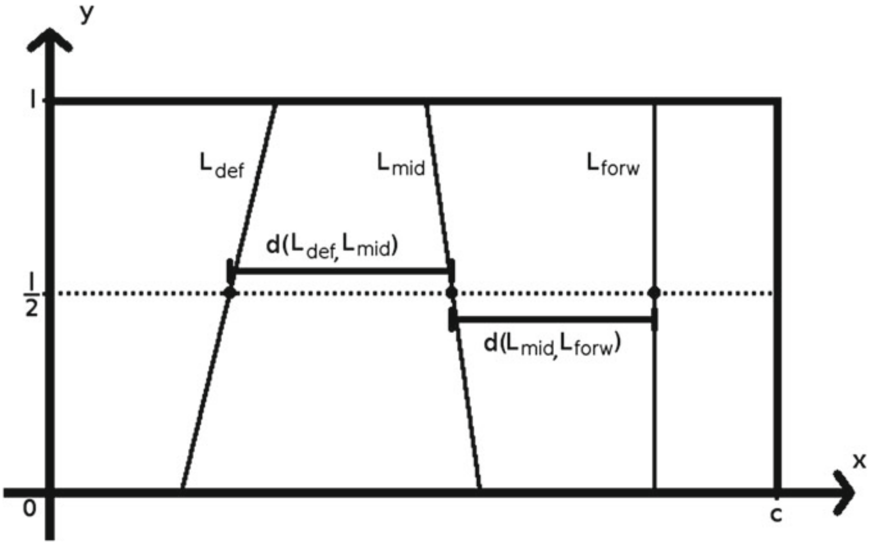

Fig. 6.4 Representation of how the distance is calculated between *Sectorial Lines* using the points of the *line* in $y = \frac{l}{2}$, where l is the length of the field

Definition 6.10 [16] Given a set of points P with N elements (each player performing the same role as the line being calculated), where element i has coordinates (x_i, y_i), the simple linear regression used to estimate the equation $y = \alpha + \beta x$ calculates α and β using the following equations:

$$\beta = \frac{\sum_{i=1}^{N}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{N}(x_i - \overline{x})^2} \tag{6.10}$$

$$\alpha = \overline{y} - \beta\overline{x}, \tag{6.11}$$

where $\overline{x}$ is the average of the values of x, and similarly $\overline{y}$ is the average of the values of y.

6.5.2 Real Life Examples

The distance between two Sectorial Lines was calculated by the differences of value of x in the y position corresponding to the center of the field, as illustrated in Fig. 6.4.

The results obtained by Team A in the match are presented in Table 6.9 with intervals of 30 s represented in Table 6.10.

No results are presented for the teams in the SSG because this metric required the existence of a fixed role for each player, which is a flexible feature in SSGs.

Table 6.9 Values obtained for the distances in the x axis between sectorial lines of Team A in $y = \frac{l}{2}$ in a match, calculated as represented in Fig. 6.4, for periods of 30 s

Period of time (s)	Distance between sectorial lines (m)		
	d(L_{def}, L_{mid})	d(L_{def}, L_{forw})	d(L_{mid}, L_{forw})
[0; 30[	39.7718	54.0001	14.2283
[30; 60[	110.9706	111.8264	20.4993
[60; 90[	29.3031	24.0419	24.7967
[90; 120[	15.0799	11.2508	12.6522
[120; 150[	314.4038	387.3574	79.8965
[150; 180]	25.9555	30.7293	9.2692

Table 6.10 Values obtained for the distances in the x axis between sectorial lines of Team A in $y = \frac{l}{2}$ in a match, calculated as represented in Fig. 6.4, in the entire period of time of 3 min

Period of time (s)	Distance between sectorial lines (m)		
	d(L_{def}, L_{mid})	d(L_{def}, L_{forw})	d(L_{mid}, L_{forw})
[0; 180]	89.4588	103.4260	26.9359

6.5.3 General Interpretation

The Sectorial Lines define the three main lines of a team [5]: (i) defensive; (ii) middle; and (iii) forward. This measure allows to identify how close are the sectors of the team and how coordinate they are. The original study verified that small values of coordination were found between Sectorial Lines of the team [5].

This measure represents one of the main observations made by coaches in top view analysis: the distance between 'lines'. The lines are the lateral displacements of the players from a sector (defensive, middle or forward). During defensive moments, the lines must be close to avoid significant spaces between them. Great spaces between lines may allow the opponent team to penetrate with the ball and exploit the free space to move forward.

Small values of distance between Sectorial Lines are expected in defensive moments. On the other hand, the attacking moments will lead to a greater dispersion mainly between forward and middle line. Coaches can use the information of the distance between lines to characterize the defensive efficacy of the team during matches and to develop tasks that promote similar spaces in training sessions.

6.6 Principal Axes of the Team

6.6.1 Basic Concepts

From the position data of a team's players, the point cloud formed by the players' positions in a time instant has two principal axes that can be calculated from the eigenvectors of the point cloud. These axes can be calculated for the entire team, or for each sector separately, using only position data of the players of a single sector.

Preposition 6.1 [1, 6] *Given a symmetric real matrix, $A \in \mathbb{R}^{n\times n}$:*

1. *all eigenvalues of A are real;*
2. *all eigenvectors of A are real;*
3. *if all eigenvalues of A are distinct, then their eigenvectors are orthogonal.*

Definition 6.11 [1] Let A be a matrix of order n. The vector $u \in \mathbb{R}^n \backslash \{0\}$ is an eigenvector of A if there exists a scalar λ such that:

$$Au = \lambda u. \tag{6.12}$$

Then, λ is an eigenvalue of A associated to the eigenvector u.

Definition 6.12 [2, 8, 13, 15, 28] Given a time-series of length N containing the positions of each player on every measured instant, the variance-covariance matrix of the data, $M \in \mathbb{R}^{2\times 2}$, can be described as follows:

$$M = \begin{bmatrix} var(x) & cov(x, y) \\ cov(x, y) & var(y) \end{bmatrix}, \tag{6.13}$$

where $var(x) = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2$ and $cov(x, y) = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})(y_i - \overline{y})$. If the nonzero eigenvalues of M are distinct, then the orthogonal eigenvectors of M define the direction of the Principal Axes of the team, and the length of the Principal Axes is given by the following equation:

$$L_i = \sqrt{\lambda_i} \times ||u_i||, i = 1, 2. \tag{6.14}$$

Remark 6.3 In this case, the eigenvalues of $M \in \mathbb{R}^{2\times 2}$ are found through the following equation:

$$det(M - \lambda I_2) = 0, \tag{6.15}$$

where the $\lambda \in R$ are the eigenvalues of M and $I_2 \in \mathbb{R}^{2\times 2}$ is the identity matrix of order 2.

From each different nonzero eigenvalue, $\lambda_i \in R$, $i = 1, 2$, of M, the corresponding eigenvector, $u_i \in \mathbb{R}^2$, $i = 1, 2$ is given as the solution of the following equation:

$$Mu_i = \lambda_i u_i. \tag{6.16}$$

Remark 6.4 The distance between the center of the Principal axes of two teams can be calculated as the Inter-team Distance, presented in Sect. 4.2.

6.6.2 Real Life Examples

A screenshot of a representation of the Inter-axes captured from the uPATO software is displayed in Fig. 6.5.

6.6.3 General Interpretation

A set of dots can have a center of gravity and its two Principal Axes [13]. In the original articles that proposed this measure, the aim was to classify the positioning of defense considering the opponent's team [13, 18]. The oscillation of both axes may represent the notion of 'block' or 'in pursuit' for the defense [12, 18]. The orientation of the axes indicates the direction of greater variance of position of the players of the team. This is better exemplified in Fig. 6.6, where the two axes point towards the direction of greater variance of the data. the length of the axes is double the length of the component vector calculated through 6.14, with L being the length of each extremity to the intersection point.

In the case of no classification of attack or defense, this measure can be readjusted to classify the interactions between Sectorial Lines.

Fig. 6.5 Screenshot of the uPATO game animation with the representation and values of the inter-axes visible for both teams in the example SSG

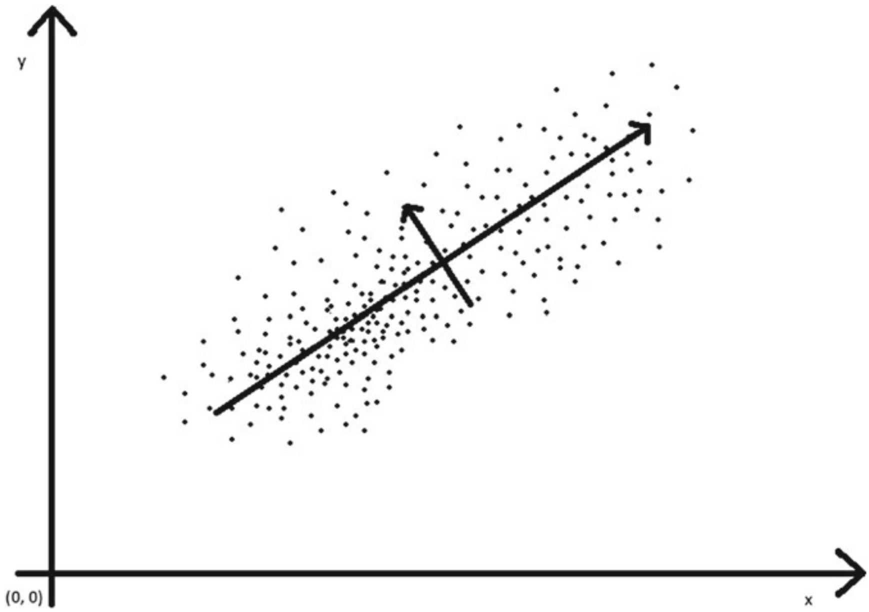

Fig. 6.6 Example scheme of the Principal Axes of a point cloud. The largest axis points towards the direction of the greater dispersion of data, while the second axis is orthogonal to the first, and points towards the second largest dispersion

Moreover, Principal Axes can be also analyzed by sectors. It can be possible to identify the variation of distances between the intersection of the defensive and attacking axes of both teams or between middle axes. This may provide useful information to identify the behavior of the lines in specific situations. Some teams will opt to approximate the forward line to the opponent's defensive line, while others will try to be far to gain some space. The relationship between middle axes will also be important to understand the dynamic in attacking and defensive moments.

6.7 Dominant Region

6.7.1 Basic Concepts

The playing field is divided in $length \times width$ squares, where each square has 1m^2 of area. Each square is attributed to the player with the least euclidean distance to it. The set of squares belonging to a player define their Dominant Region.

Definition 6.13 [24] Given a time-series of length N containing the positions of each player on every measured instant, and dividing the field of play in $length \times width$

squares, the Dominant Region of a player p on instant t is given by the following equation:

$$DR_p(t) = \left\{ s_{i,j} \in s : min\Big(d(s_{i,j}, k)\Big); k = p, k = 1, 2..., N_{players} \right\}, \tag{6.17}$$

in which $d(s_{i,j}, k) = \sqrt{(x_{s_{i,j}} - pos_{k_x}(t))^2 + (y_{s_{i,j}} - pos_{k_y}(t))^2}$, and where s is the set of all squares that make up the field of play, $s_{i,j}$ the square on position (i, j) of the field, and (pos_{k_x}, pos_{k_y}) the position of player k.

Definition 6.14 [24] Given a set of N squares, defining the Dominant Region of a player, the area of the Dominant Region is given by:

$$DM_{Area} = \sum_{i=1}^{N} s_i. \tag{6.18}$$

Definition 6.15 [24] Given a set of Dominant Region areas, DM_a, containing the different Dominant Region areas of each player of a team, the area of the Dominant Region of a team is given by:

$$DM_{team} = \sum_{i=1}^{N_{players}} DM_{a_i}. \tag{6.19}$$

6.7.2 Real Life Examples

The results obtained by both teams in the SSG are represented for the entire 3 min in Fig. 6.7.

A screenshot of a representation of the Dominant Region captured from the uPATO software is displayed in Fig. 6.8.

No results are presented for the match because this metric required the existence of data from both teams, which is not available in the evaluated match.

6.7.3 General Interpretation

The Voronoi region can be used to classify the spatial territory of a player [29]. This measure allows to quantify the spatial partitioning of the pitch area into cells, each associated with players according to their positions [11]. The cells result from applying the concept of nearest-neighbor rule in which each player is associated to all parts of the pitch that are nearer to that player than to any other player [11,

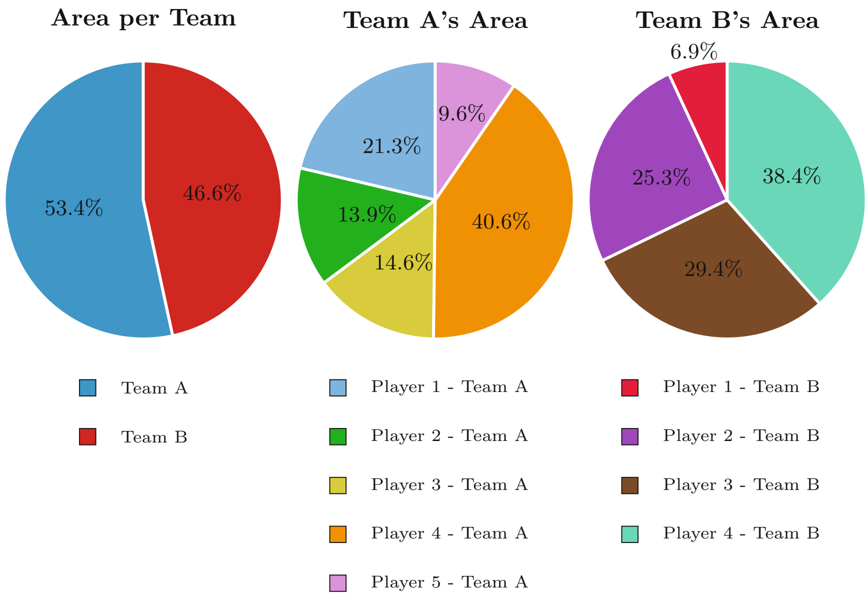

Fig. 6.7 Representation of the values obtained in the dominant region visible in the example SSG

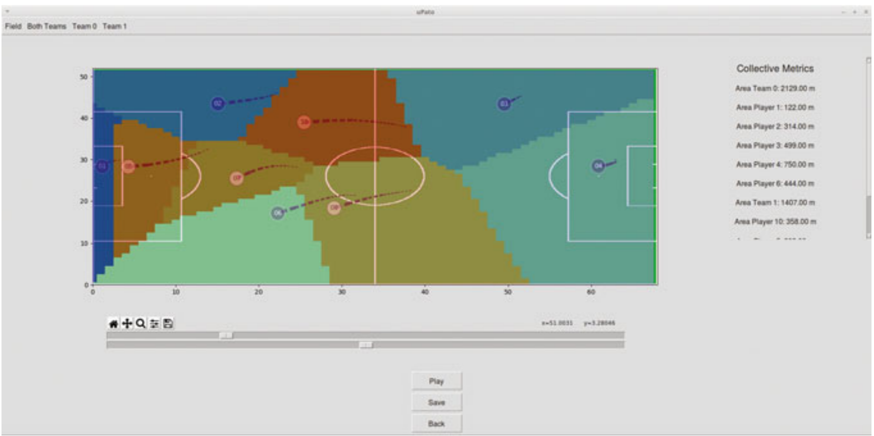

Fig. 6.8 Screenshot of the uPATO game animation with the representation and values of the dominant region visible in the example SSG

22]. Voronoi diagrams may help describe the interaction between the two teams by comparing the spatial pattern formed by the players and the oscillations of spatial occupation of players in specific moments of the game [10].

Moreover, the observation of individual spatial regions can characterize the capacity and dominance of players during the match and quantify the zones of influence of each player. The collective measure may also determine the territorial influence of the team during the game and to monitor which zones are more controlled by a team. This measure can help to classify the patterns of spatial occupation of teams and more specifically of players.

6.8 Major Ranges

6.8.1 Basic Concepts

The Major Range of each player is defined as an ellipse centered on its average position, with its axes defined as the standard deviation of the player's movement.

Definition 6.16 [31] Given a time-series of length N containing the positions of each player on every measured instant, the Major Ranges of the team are the ellipses defined, for each player of the team, by the following equation:

$$\frac{(x-\overline{x})^2}{\sigma_x^2}+\frac{(y-\overline{y})^2}{\sigma_y^2}=1, \tag{6.20}$$

where $(\overline{x}, \overline{y})$ is the average position of a player during the time-series and σ represents the standard deviation of the player's position during the time series.

6.8.2 Real Life Examples

The results obtained by a player in each team in the SSG are presented for periods of 30 s in Table 6.11 and for the entire 3 min in Table 6.12.

Table 6.11 Values obtained for the Major Ranges of a player from each team in x in an SSG, for periods of 30 s

Period of time (s)	Team A		Team B	
	Avg. position (m)	Std. deviation (m)	Avg. position (m)	Std. deviation (m)
[0; 30[	20.9644	5.3399	33.5689	2.2508
[30; 60[	13.8660	6.4446	27.1063	12.4493
[60; 90[	17.0228	2.0186	39.2959	8.3456
[90; 120[	23.0112	4.8530	33.1544	4.0458
[120; 150[	15.5190	6.9755	36.0338	4.9149
[150; 180]	22.3510	4.8749	32.6194	1.7885

These can be compared to those obtained by two players in the match, presented in Table 6.13 with intervals of 30 s and for the entire 3 min in Table 6.14.

A screenshot of a representation of the Major Ranges captured from the uPATO software is displayed in Fig. 6.9.

Table 6.12 Values obtained for the Major Ranges of a player from each team in x in an SSG, in the entire period of time of 3 min

Period of time (s)	Team A		Team B	
	Avg. position (m)	Std. deviation (m)	Avg. position (m)	Std. deviation (m)
[0; 180]	18.7867	6.3735	33.6169	7.6923

Table 6.13 Values obtained for the Major Ranges of two players from Team A in x in a match, for periods of 30 s

Period of time (s)	Player 1		Player 2	
	Avg. position (m)	Std. deviation (m)	Avg. position (m)	Std. deviation (m)
[0; 30[	74.0947	2.5430	90.4967	4.1956
[30; 60[	67.1143	4.7815	83.4681	0.7011
[60; 90[	68.4616	8.7448	90.9741	6.6659
[90; 120[	91.4237	3.1986	98.1573	2.2631
[120; 150[	65.0343	6.8399	81.1200	4.3843
[150; 180]	81.9761	2.4285	91.6101	4.7786

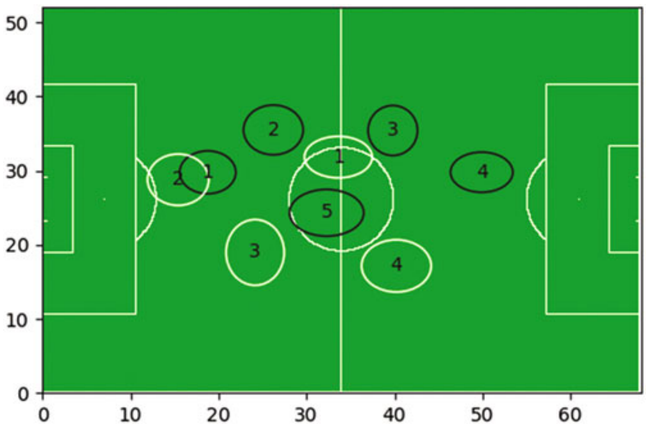

Fig. 6.9 Screenshot of the uPATO collective metrics representation of the Major Ranges in the example SSG

Table 6.14 Values obtained for the Major Ranges of two players from Team A in x in a match, in the entire period of time of 3 min

Period of time (s)	Player 1		Player 2	
	Avg. position (m)	Std. deviation (m)	Avg. position (m)	Std. deviation (m)
[0; 180]	74.6838	10.7423	89.3115	7.0518

6.8.3 General Interpretation

The Major Ranges approach was introduced in a soccer case study [31]. This measure contributes to assess the division of labor between players in a team [7]. As described by Duarte et al. [7] "the predominant area of each individual's interventions during performance is defined by an ellipse centered at the 2-dimensional mean location of each performer, with semi-axes being the standard deviations in X and Y directions, respectively".

This measure represents the range of a player during different periods of the match and may provide important information about variations of spatial occupation in different scenarios. It can be also assessed the coordination between teammates during the performance and identify if the patterns of spread or contract may be associated between playing roles. Task constraints (e.g., opponents, goal, match status, possession of the ball) may influence the variations of Major Ranges in the players. The variation of the ellipses in longitudinal or lateral axes may also indicate some patterns to explore different playing styles in defensive and attacking phases. In attacking moments, an ellipse with prominence in longitudinal axis may indicate a tendency to exploit the direct playing style. In the other hand, a higher range in lateral axis may indicate that the team varies the zone of play by using circulation of the ball. The variation of the ellipse in different periods of the match can be also useful to identify how team's behave over the match in their playing style.

6.9 Identify Team's Formations

6.9.1 Basic Concepts

From the position data of a team's players, and taking into account the roles defined for each player initially, by taking the average position of a player on a given role, and forming a cost matrix based on the distance of a player to a role, the Hungarian Method can be applied to find the minimum cost for attributing each role to a player, based on its current position.

Definition 6.17 [3] Given a time-series of length N containing the positions of each player on every measured instant, the cost matrix of a player occupying each role, in each instant, can be calculated as follows:

$$CM = \begin{bmatrix} d(\overline{pos_1}, pos_{p_1}) & \dots & d(\overline{pos_1}, pos_{p_k}) \\ \vdots & \ddots & \vdots \\ d(\overline{pos_i}, pos_{p_1}) & \dots & d(\overline{pos_i}, pos_{p_k}) \end{bmatrix}, \tag{6.21}$$

where $d(p, q)$ represents the euclidean distance between two points, p and q, and $\overline{pos_i}$ represents the average position of the player associated with role i, and pos_{p_k} represents the position of player k in an instant.

Definition 6.18 [3, 17] Given a cost matrix CM, the minimum cost for each player occupying a specific position p can be given by the following algorithm:

Algorithm 3: Hungarian Method to find the minimum cost for assigning of each player to each position.

1 Subtract the smallest entry in each row from all the entries of its row.
2 Subtract the smallest entry in each column from all the entries of its column.
3 Draw lines through appropriate rows and columns so that all the zero entries of the cost matrix are covered and the minimum number of such lines is used.
4 *Test for Optimality:*(i) If the minimum number of covering lines is n, an optimal assignment of zeros is possible and the algorithm is finished. (ii) If the minimum number of covering lines is less than n, an optimal assignment of zeros is not yet possible. In that case, proceed to Step 5.
5 Determine the smallest entry not covered by any line. Subtract this entry from each uncovered row, and then add it to each covered column. Return to Step 3.

6.9.2 Real Life Examples

The results obtained by Team A in the match are presented in Table 6.15 with intervals of 0.1 s.

6.9.3 General Interpretation

The aim of team's formation is to classify the regular playing position of players in specific periods or moments of the match [14]. An approach on soccer preprocessed the trajectories of players and segmented the positions into game phases [3, 30]. Another approach was made in field hockey [20].

This measure can be used to visualize the most recurrent position of players in periods of the match and based on the status of possession. Moreover, coaches can

Table 6.15 Values obtained for the team formation from Team A in a match, in each instant for 1 s

Period of time (s)	Current position								
	Player 1	Player 2	Player 3	Player 4	Player 5	Player 6	Player 7	Player 8	Player 9
[2.0; 2.1[	1	2	6	4	5	3	7	8	9
[2.1; 2.2[	1	2	6	4	5	3	7	8	9
[2.2; 2.3[	1	2	3	4	5	6	7	8	9
[2.3; 2.4[	1	2	3	4	5	6	7	8	9
[2.4; 2.5[	1	2	3	4	5	6	7	8	9
[2.5; 2.6[	1	2	3	4	5	6	7	8	9
[2.6; 2.7[	1	2	3	4	5	6	7	8	9
[2.7; 2.8[	1	2	3	4	5	6	7	8	9
[2.8; 2.9[	1	2	3	4	5	6	7	8	9
[2.9; 3.0]	1	2	3	4	5	6	7	8	9

use this information to classify the formation of the team and the variations that emerge during the match. The opponents can be also classified by this measure, thus providing information about the spatial mean spatial territory of players and the numeric relationship by sectors. However, it is important to highlight that this measure only represents a static position and other dynamic measures must be used to identify the territory and the typical movements of the players.

References

1. Alexander Ed, Poularikas, D (1998) The handbook of formulas and tables for signal processing. CRC Press, Boca Raton, FL, USA, pp 73–79
2. Beezer RA (2008) A first course in linear algebra. Beezer
3. Bialkowski A, Lucey P, Carr P, Yue Y, Matthews I (2014) Win at home and draw away: automatic formation analysis highlighting the differences in home and away team behaviors. In: Proceedings of 8th annual MIT sloan sports analytics conference. pp 1–7
4. Cizek V (1970) Discrete hilbert transform. IEEE Trans Audio Electroacoust 18(4):340–343
5. Clemente FM, Couceiro MS, Martins FML, Mendes RS, Figueiredo AJ (2014) Developing a football tactical metric to estimate the sectorial lines: a case study. Computational science and its applications. Springer, Berlin, pp 743–753
6. Cullum JK, Willoughby, RA (2002) Lanczos algorithms for large symmetric eigenvalue computations, vol. 1: Theory. SIAM
7. Duarte R, Araújo D, Correia V, Davids K (2012) Sports teams as superorganisms: implications of sociobiological models of behaviour for research and practice in team sports performance analysis. Sport Med 42(8):633–642
8. Euler L (1767) Du mouvement d'un corps solide quelconque lorsqu'il tourne autour d'un axe mobile. Mémoires de l'Académie des Sciences de Berlin, 16(1760):176–227
9. 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)

10. Fonseca S, Milho J, Travassos B, Araújo D, Lopes A (2013) Measuring spatial interaction behavior in team sports using superimposed voronoi diagrams. Int J Perform Anal Sport 13(1):179–189
11. Fonseca S, Milho J, Travassos B, Araújo D (2012) Spatial dynamics of team sports exposed by voronoi diagrams. Hum Mov Sci 31(6):1652–1659
12. Grehaigne JF, Bouthier D, David B (1997) Dynamic-system analysis of opponent relationships in collective actions in soccer. J Sports Sci 15(2):137–149 PMID: 9258844
13. Gréhaigne JF, Mahut B, Fernandez A (2001) Qualitative observation tools to analyse soccer. Int J Perform Anal Sport 1(1):52–61
14. Gudmundsson J, Horton M (2017) Spatio-temporal analysis of team sports. CSUR 50:1–34
15. Jolliffe I (2002) Principal component analysis. Wiley Online Library
16. Kenney JF, Keeping ES (1962) Linear regression and correlation. Mathematics of statistics. Van Nostrand, Princeton, NJ, pp 252–285
17. Kuhn HW (1955) The Hungarian method for the assignment problem. Nav Res Logist Q 2(1–2):83–97
18. Lemoine A, Jullien H, Ahmaidi S (2005) Technical and tactical analysis of one-touch playing in soccer–study of the production of information. Int J Perform Anal Sport 5(1)
19. Likas A, Vlassis N, Verbeek JJ (2003) The global k-means clustering algorithm. Pattern Recognit 36(2):451–461
20. Lucey P, Bialkowski A, Carr P, Morgan S, Matthews I, Sheikh Y (2013) Representing and discovering adversarial team behaviors using player roles. In: 2013 IEEE conference on computer vision and pattern recognition. pp 2706–2713, June
21. Nagy M, Akos Z, Biro D, Vicsek T (2010) Hierarchical group dynamics in pigeon flocks. Nature 464(7290):890–893
22. Okabe A, Boots B, Sugihara K, Chiu SN (2000) Spatial tesselations: concepts and applications of Voronoi diagrams. Wiley, New York
23. Palut Y, Zanone P-G (2005) A dynamical analysis of tennis: concepts and data. J Sports Sci 23(10):1021–1032
24. Rein R, Raabe D, Perl J, Memmert D (2016) Evaluation of changes in space control due to passing behavior in elite soccer using voronoi-cells. In: Proceedings of the 10th international symposium on computer science in sports (ISCSS)
25. Ric A, Hristovski R, Gonçalves B, Torres L, Sampaio J, Torrents C (2016) Timescales for exploratory tactical behaviour in football small-sided games. J Sports Sci 34(18):1723–1730 PMID: 26758958
26. Ric A, Hristovski R, Gonçalves B, Torres L, Sampaio J, Torrents C (2016) Timescales for exploratory tactical behaviour in football small-sided games. J Sports Sci 34(18):1723–1730
27. Silva P, Vilar L, Davids K, Araújo D, Garganta J (2016) Sports teams as complex adaptive systems: manipulating player numbers shapes behaviours during football small-sided games. SpringerPlus 5(1):191
28. Taboga M (2012) Lectures on probability theory and mathematical statistics. CreateSpace Independent Pub
29. Taki T, Hasegawa J (2000) Visualization of dominant region in team games and its application to teamwork analysis. In: Proceedings of the international conference on computer graphics, CGI ’00. IEEE Computer Society, Washington, DC, USA, pp 227–235
30. Wei X, Sha L, Lucey P, Morgan S, Sridharan S (2013) Large-scale analysis of formations in soccer. In: 2013 international conference on digital image computing: techniques and applications (DICTA), pp 1–8, Nov
31. Yue Z, Broich H, Seifriz F, Mester J (2008) Mathematical analysis of a football game. part I: individual and collective behaviors. Stud Appl Math 121(3):223–243

Appendix
Available Metrics in uPATO

As described in Sect. 2.2.2, uPATO divides metrics into individual and collective metrics. This appendix contains a full list of all the metrics included in each type of metrics.

The set of individual metrics includes:

- Kolmogorov Entropy;
- Shannon Entropy;
- Distances performed by speed (walk, jog, run and sprint);
- Total distance;
- Sprint volume;
- Maximum speed;
- Spatial Exploration Index;
- Displacement Angle.

And the set of collective metrics includes:

- Geometrical Center;
- Stretch Index;
- Surface Area;
- Team Width and Length;
- lpwratio;
- Team Separateness;
- Inter-team distance;
- Time Delay;
- Coupling Strength;
- Principal Axes;
- Directional Correlation Delay;
- Intra-team Coordination Tendencies;
- Sectorial Lines;
- Inter-player Context;
- Dominant Region.

F.M. Clemente et al., *Computational Metrics for Soccer Analysis*, SpringerBriefs in Applied Sciences and Technology, https://doi.org/10.1007/978-3-319-59029-5