Current Approaches to Tactical Performance Analyses in Soccer Using Position Data

Daniel Memmert¹ · Koen A. P. M. Lemmink² · Jaime Sampaio³

Abstract Tactical match performance depends on the quality of actions of individual players or teams in space and time during match-play in order to be successful. Technological innovations have led to new possibilities to capture accurate spatio-temporal information of all players and unravel the dynamics and complexity of soccer matches. The main aim of this article is to give an overview of the current state of development of the analysis of position data in soccer. Based on the same single set of position data of a high-level 11 versus 11 match (Bayern Munich against FC Barcelona) three different promising approaches from the perspective of dynamic systems and neural networks will be presented: Tactical performance analysis revealed inter-player coordination, inter-team and inter-line coordination before critical events, as well as team-team interaction and compactness coefficients. This could lead to a multi-disciplinary discussion on match analyses in sport science and new avenues for theoretical and practical implications in soccer.

Key Points

Position data can be used to develop collective performance indicators capable of describing and understanding the dynamics of the match. Processing measures such as the distance to positional-centroid and speed facilitate assessment of inter-player coordination and depict different predictabilities in players’ movement behaviour that ultimately contribute to the overall team organization.

Position data can be used to explore the dynamic patterns before critical events such as goal-scoring opportunities and goals. Variability of distances between centroids of teams and lines can be used to shed light on inter-team and inter-line coordination in these critical periods and might help to understand the process of scoring goals or to prevent this.

With a hierarchy of several artificial neural networks, it is possible to automatically classify complex and characteristic tactical patterns, formations of tactical groups and their dynamic changes (e.g. compactness) and interactions on the basis of position data. ‘Compactness’ is defined as the movement speed of all players targeted at recovering ball possession as fast as possible.

1 Introduction

While there do exist review articles about the assessment of technical skills [1], the assessment of tactical behaviour in team sports has not been paid a lot of attention until a
a couple of years ago [2–7]. State-of-the-art research as well as public interest is calling for a detailed and objective scientific analysis of soccer matches [8–11]. The primary aim of this leading article is the discussion of some novel objective analysis tools to evaluate tactical performance patterns in high-level men’s soccer. Therefore, we present and discuss three different recent approaches based on one single set of position data, namely the Champions League Quarterfinal of Bayern Munich against FC Barcelona from the 2008/2009 season (14 April 2009).

The traditional and present standard is to assess tactical performance during game play by means of the method of game observation [12]. Qualitative game observations are less objective and systematic (e.g. structured and comprehensive), use subjective impressions of the observers, are very slow in analysis processes (4–6 h), and take advantage of the experiences and ‘know-how’ of experts. In contrast, quantitative game observations proceed very objectively by using predefined observation schedules (category systems) to collect the data of game behaviour [13]. Subsequently, these data are evaluated and indices are calculated to value the whole player performances or individual performance components [6, 14, 15]. Just over the past years, progress of computer science made it possible to provide position data and thus track the players’ movements [16–18]. Position tracking systems determine the positions of the 22 players and the ball in x- (parallel to sidelines) and y-coordinates at 25 frames per second, which amounts to approximately 135,000 positions per object and match, and a total of about 3,100,000 positions for all players and the ball. There will be a dramatic speed advantage concerning the evaluation of the position data (reducing from 6–8 h to several minutes). The small effort for data acquisition will enable the accumulation of enormous amounts of data and thus bring new opportunities for theory construction and practice in sports.

While physical parameters such as distance covered or speed have been automatically analysed for a few years, there are only few approaches doing this with tactical parameters in team sports (see Table 1). In this article we introduce and discuss three current approaches: measuring inter-player coordination [30, 33], measuring inter-team coordination before critical events [34, 35] and measuring team-team interaction and compactness coefficients [36, 37]—all of which are presented in the light of the results of a single set of position data from a high-level 11 versus 11 match (Bayern Munich against FC Barcelona).

2 Tactical Performance Analysis in Soccer by Measuring Inter-Player Coordination

Match performance in soccer results from the players’ individual and multidimensional skills, expressed within a collective organisation. These tactical organisations mainly refer to how teams are distributed in the pitch across time and, anecdotally, are held responsible either for increasing or decreasing the overall productivity of the players. Therefore, the position data can be used to develop collective performance indicators capable of describing and understanding the dynamics of these complex, non-linear and chaotic organisations.

<table>
<thead>
<tr>
<th>Key performance index</th>
<th>Method</th>
<th>Description</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length, width, space</td>
<td>Distance</td>
<td>Measures the average expansion of a team in the direction of x and y or rather in both dimensions</td>
<td>Castellano et al. [19], Moura et al. [20]</td>
</tr>
<tr>
<td>Space control</td>
<td>Voronoi</td>
<td>Models space control with the help of a Voronoi diagram</td>
<td>Fujimura and Sugihara [21], Fonseca et al. [22], Taki and Hasegawa [23], Kang et al. [24], Horton et al. [25]</td>
</tr>
<tr>
<td>Event recognition</td>
<td>Rule-based, decision trees</td>
<td>Recognizes events from position data such as passes, goals, offside; rule-based system</td>
<td>Gudmundsson and Wolle [26], Wei et al. [27]</td>
</tr>
<tr>
<td>Route clustering</td>
<td>Clustering algorithms (Fréchet distance)</td>
<td>Filters subgroups from movement patterns of one or more players</td>
<td>Gudmundsson and Wolle [26], Hirano and Tsumoto [28]</td>
</tr>
<tr>
<td>Pass evaluation</td>
<td>Motion model and passable area</td>
<td>Calculates regions, such as the ‘passable area’, for every pass based on a motion model and evaluates passes according to difficulty or decision quality</td>
<td>Horton et al. [25], Gudmundsson and Wolle [29]</td>
</tr>
<tr>
<td>Distance from the team’s centre</td>
<td>Euclidian metrics</td>
<td>Calculates the players’ average, minimal and maximal distance from the team’s centre</td>
<td>Sampaio and Macães [30], Białkowski et al. [31]</td>
</tr>
<tr>
<td>Formation</td>
<td>Mean average, main component analysis</td>
<td>Calculates the average position and thus determines an actual tactical formation</td>
<td>Białkowski et al. [32]</td>
</tr>
</tbody>
</table>
Current studies have already identified several measures that may be candidates for developing key collective variables, such as the team centroid, stretch index or surface area [30]. The team centroid, calculated as the mean position of all outfield players from one team, exhibits low variability when used to measure inter-player coordination at the macro level (11-a-side match). However, the interaction between each player and his position-specific centroid (e.g. defenders, midfielders and forwards) has been identified as a strong candidate to capture the players’ movement behaviour more accurately [30]. The data processing techniques that can fit the purposes of identifying these interactions can include different approaches such as approximate entropy (ApEn) and relative phase. The approximate entropy values are used to identify predictability in players’ movement patterns, understood as non-linear time series that incorporate both deterministic chaotic and stochastic processes [38]. The algorithm quantifies predictability in a time series, by measuring the logarithmic likelihood of patterns that are close for m consecutive observations and remain close even for $m + 1$

![Fig. 1 Team formation and predictability of distances between players and their positional-centroid. The results show the players’ predictability in relation to their positional-centroid, as measured by the approximate entropy from the distance between each player and their positional-centroid. A cluster analysis automatically classified the predictabilities in three different groups (higher, medium and lower). The performance of the highly predictable players is also described with the average ± standard deviation of the distance to positional-centroid. GK goalkeeper, CD central defender, LD left defender, RD right defender, CM central midfielder, LM left midfielder, RM right midfielder, FW forward, CF central forward, RF right forward, LF left forward, Cdef defensive centroid, Cmid midfield centroid, Cfw forward centroid.](image)
as highly predictable, probably reflecting his role as a strong communicator between defenders and attackers. Further developments of this analysis can include examining how these interactions change from the first to the second half of the match due to increasing fatigue, strategic or situational changes. In addition, coaching staffs can also test the effects of changing team formations (using different distributions of players in defence, midfield or offence sectors, e.g., 4:3:3, 4:4:2, 4:5:1) and using different player combinations for each specific position or sector.

Relative phase procedures can also be used to identify inter-player coordination. This is a non-linear signal processing descriptor for synchronisation between two oscillators [40]. The relative phase calculation is often carried out by using a Hilbert transform [41] and the obtained values are expressed in angles. The values close to 0° refer to simultaneous patterns of synchronisation, often referred to as in-phase, whereas values close to 180° refer to asynchronous patterns of synchronisation, often referred to as moving in anti-phase [42].

As an example, Fig. 2 shows three different confrontations that occurred in the matches, namely, defenders versus opponents’ forwards, midfielders versus midfielders and forwards versus opponents’ defenders. The results (relative phase results) and the percentage of near-in-phase represents the amount of time that players were synchronised in each half of the match. The values close to 0° and 360° multiples refer to simultaneous patterns of synchronisation (i.e. in-phase, in which speeds are synchronised and both are increasing or both are decreasing); whereas values close to 180° and 360° multiples refers to asynchronous patterns of synchronisation (i.e. anti-phase, in which speeds are synchronised and present different trends, one increasing and the other one decreasing or vice-versa).
show the average speed of the players, which constitute each positional centroid (Fig. 2a, c, e), and how these different speeds are synchronised during match time (Fig. 2b, d, f, relative phase results). The results display how Bayern Munich defenders were faster than FC Barcelona forwards, particularly in the second half (+8.6 and +41.3 %, respectively). The FC Barcelona midfielders were faster than their counterparts for a substantial amount of time (+53.2 and +54.6 %, respectively) and the Bayern Munich forwards were faster than FC Barcelona defenders (+18.9 and +18.4 %, respectively). Figure 2b, d, f show that players were highly synchronised in all situations during the initial match periods. The second half was played with less near in-phase synchronisation [42], particularly for the midfielder confrontations (Fig. 2c, d).

3 Tactical Performance Analysis in Soccer by Measuring Inter-Team and Inter-Line Coordination Before Critical Events

To explore the tactical behaviour of interacting players or teams, there is a need for collective variables that capture the dynamics of this behaviour. Centroids, stretch indexes, length and widths, and surface areas of teams seem to provide a sound basis for collective variables that capture the dynamics of attacking and defending in soccer at team level [10]. The potential of centroids to describe the rhythmic flow of attacking and defending has been shown in small-sided games and 11-a-side soccer matches [2, 14, 34, 35]. These findings are in line with the general aim of the game, as such patterns reflect the two teams moving up and down the field to arrive in a scoring position or prevent that from happening.

From a dynamic system perspective, it is expected that high variability in inter-team distance (distance between two teams’ centroids) reflects perturbations in the balance between the teams’ behaviour that precede critical game events such as goal attempts or goals. However, in an 11-a-side soccer match periods of high variability were associated with collective defensive actions and team reorganisation in dead-ball moments rather than goals or goal attempts [35]. This may be explained by the fact that different functional units exist within and between teams and that teams switch between these functional states depending on the situational context to preserve balance between them. From this perspective variability in inter-line distances, i.e. distances between centroids of attackers, midfielders and defenders of both teams before critical events, could shed light on this issue.

The match Bayern Munich against FC Barcelona was first scanned for all critical events, meaning goal attempts and goals. The 30 s prior to a critical event were defined as the critical period. Based on team and line centroids inter-team and inter-line distances were calculated. Variability was determined as the rate-of-change of these variables. Differences between variability of inter-team distances and inter-line distances were determined with an independent sample t test ($p < 0.05$). Based on a method described in an earlier study [35], the threshold for high variability was defined as three times the standard deviation of the total match variability. Additionally, the strengths of the different couplings and frequencies of variability above the threshold were calculated.

In general, results clearly show that soccer is an in-phase sport with high couplings between teams and lines especially in longitudinal direction, indicated by the explained variance between team and line centroids ($0.86 < R^2 < 0.99$). In total, nine critical events were identified, seven goal attempts and two goals. In Fig. 3 the variability of the inter-line distance between the attacking line and defending line of the opponent before a goal is presented. The centroid of the attacking line is crossing the centroid of the defending line (Fig. 3a) and variability rises above the inter-line threshold between 20 and 10 s before the goal is scored (Fig. 3c). This moment in time matches the moment of crossing of the attacking and defending line (Fig. 3).

Differences in variability during the 30 s prior to a goal attempt or goal were significantly higher for inter-line distances than for inter-team distances ($p < 0.05$). As an example, in Fig. 4 the variability of the inter-line distance is especially higher in the period between 20 and 10 s before the goal compared to the inter-team distance. These differences were seen in eight out of nine critical periods but at different moments in the 30-s periods.

On average, variability in the 30-s periods before a critical event was not higher than the variability of a random 30-s game period. However, when focusing on the variability of the separate 10-s periods before a critical event some interesting findings were seen. On average, variability seemed to increase toward the critical moments. However, there were different patterns for the separate lines. For the defending line (defenders of the attacking team vs. the attackers of the defending team) and the midfield line (midfielders vs. midfielders), no clear increase or decrease of variability of inter-line distances was seen before the critical moments. In contrast, the attacking line (attacking line (attackers of the attacking team versus the defenders of the defending team), a clear increase of variability of the inter-line distance was seen approaching the critical moments. So, the variability of the attacking line of the attacking team and the defending line of the defending team seems most important before a critical event. This phenomenon was also illustrated by the higher average number of moments of high variability for this inter-line distance.
Fig. 3 Example of a critical period (30 s) before a goal: 
(a) position of the centroids in longitudinal direction of the attacking line of the scoring team and the defending line of the opposite team; 
(b) distance between the centroids of both lines (inter-line distance); 
(c) rate of change (RoC) of the inter-line distance as a measure of variability: the threshold was defined as three times the standard deviation of the match variability.
during the critical periods (4.4 for the attacking line vs. 1.4 for the midfielders and 0.2 for the defending line). Interestingly, the mean variability for the inter-team distances between midfielder centroids was the highest 30 to 20 s before the critical events. This could illustrate that a higher variability of the inter-line distance between midfielders of both teams earlier in time indicates a disturbed balance between teams that leads to an increase of variability in the attacking line (attackers of the attacking team versus the defenders of the defending team) before a critical event. Future studies should incorporate different time scales and other functional units of players in relation to tactical strategy or ball location.

4 Tactical Performance Analysis in Soccer by Measuring Team-Team Interaction and Compactness

A self-organizing map is a type of artificial neural network, which in particular can be used for pattern or type analysis [43, 44]. Simply put, a self-organizing network consists of neurons which gather information from training data: During one’s learning phase, training input stimulates and influences the ‘winner’ neuron (i.e. the most connected to the incoming information) by improving its receptiveness for different pieces of information of incoming training data.). Also, the neurons of a number of neighbour shells (i.e. those next to the winner neuron) are modified in the same way, activating a larger neuronal area with the same or lower intensity of the stimulus.

More specifically, training data contain information about the spatio-temporal formation of players on the playing field. According to a self-organizing system this information can be distributed to an artificial neural network consisting of a number of boxes, so-called ‘artificial neurons’. Each neuron represents a spatio-temporal variant of the players’ formations on the playing field [45]. If a number of neurons represents a set of similar variants of that spatio-temporal type, the neurons form clusters. The result of training thus is that classes of similar input values (e.g. player position) are mapped to clusters of neurons. This means that training input values can be classified as belonging to specific clusters and therefore the wide variety of data can be organised into a small set of characteristic types.

To analyse tactical processes in soccer according to the game’s dynamic development, it is useful to use the net approach and to restrict the formations to those of tactical groups such as offence and defence. These two groups have a smaller number of members and their data should not include their location on the playing field. The data analyses only reveal the temporal sequences of these tactical formations and represent the dynamic game process (the position information, i.e. the mean value of the players’ positions, also called ‘centroid’ is kept separately for further analyses). Hence, several specific process analyses can be carried out, two of which are briefly introduced in the following examples.

The frequencies of formations and their interactions can be measured and thus can give information about which formations are preferred by which team, and which tactical
concepts are represented by those distributions [46]. Furthermore, the tactical responses of, for example, defence of one team against offence activities of the other team can be analysed and measured. Finally, if the meaning of ‘success’ is defined, success of actions and interactions can be measured in the context of formations—or in turn, as is outlined in the following, can help to characterise a specific constellation as successful or not. For example, Team A (FC Bayern Munich) defensive formation (see Fig. 5; defensive and defensive midfield) is most successful if they interact in the defensive pattern against the most frequently employed offensive formation (formation 2) of FC Barcelona.

Constellation data can be compared with other data such as ball win, possession or loss, and therefore formations can be valued time-dependently as successful or unsuccessful. This way, a tactical interaction of formations can be said to be successful for a team if the respective tactical group of the team was successful with the activity in its constellation. The prevailing question is, which particular team formations lead to success, for example ball possessions and goals? Using these data, both the team’s player formations over space and time as well as the formations of the two teams as a pair, can be analysed. The respective results deliver a statistical success-distribution according to specific (inter)action patterns within and between the teams. For example, team ‘compactness’ (=movement speed of all players targeted at recovering ball possession as fast as possible) in soccer can be calculated by a dynamical process analysis with precise point-in-time selection as a measurement (see Fig. 6). The process velocity (in m/s) is calculated to show how quickly each player moves towards the ball to try to win it back. For example, in the first half, the centre striker of FC Bayern Munich (mean = 4.9 m/s) has the weakest values, whereas the other players of FC Bayern Munich are relatively similar. Based on the individual values of all field players, the average speed of a team, the ‘compactness’ performance parameter of the defensive team strategy, can be calculated. Here, no significant difference between FC Bayern Munich (mean = 4.5 m/s) and FC Barcelona (mean = 4.3 m/s) was found.

The qualitative analysis of interaction now can be completed for example by quantitative analysis of frequencies, for instance recognising a negative rate as an indicator of significant weakness, as insignificantly rare, or just as a creative (i.e. providing an adequate and rare solution) tactical variant in order to confuse the opponent team [36, 37]. Finally, the time series of constellation types as a whole offers much information about the dynamic interaction of opponent tactical groups, containing information about frequent stereotypical interaction patterns as

---

**Fig. 5** Example of a visual representation of an offensive formation of FC Barcelona (playing from left to right) and a defensive formation of FC Bayern Munich (playing from right to left) as a specific offensive-defensive interaction pattern. Use of such visual representations allowed the authors to determine all characteristic formations of both teams and find the most important characteristic interaction patterns between both teams. The most frequent defensive pattern of FC Bayern Munich (formation 3) was 40% successful (i.e. obtained possession of the ball) against the most frequently employed offensive formation (formation 2) of FC Barcelona.

**Fig. 6** A dynamic process analysis with precise point-in-time selection as a measurement of compactness in soccer (from left to right). The ball (circle within square) and the pathway of the ball are illustrated. Team A (black squares, FC Bayern Munich) lost the ball and all players from Team A are trying to get the ball back from FC Barcelona (grey squares) as quickly as possible. The process speed is calculated (in m/s) to show how quickly each player of each team is running towards the ball; this can be calculated as an average according to the team performance for the variable ‘compactness’.
well as about rare but striking and perhaps creative ones. In order to better detect such interaction patterns, those time series—also called trajectories—again can be taken as input data and be trained to a second level network, which then allows for recognising specific types of interaction [46].

5 Conclusion

According to prevalent belief, a high level of tactical competence is crucial for players’ effective performance in team sports. Working with position data, the discussed approaches in our leading article examine how patterns emerge from the interaction of their many degrees of freedom or constraints. For the first time, we demonstrated different kinds of computer science approaches that enable to obtain and analyse new parameters such as: inter-player coordination, inter-team coordination before critical events, and team-team interaction and compactness coefficients. All these performance analysis tools could help coaches to modify their training methods (e.g. focusing on recent trends in game philosophy and tactics) according to their needs and to improve the tactical behaviour of their players.

Also these new approaches based on position data could be an important step towards objectification of tactical performance components in team sports. This would be of great interest not only for talent selection in different kinds of youth sports (basketball, team handball, soccer, field hockey, tennis), but also for the analysis of professional team sports. In addition to providing information on sports performance and their interactions, the diagnostic possibilities of these new position data methods should be considered. This can be helpful in sports practice (e.g., for sophisticated assessments of player performances) and science (e.g., for the evaluation of sports teaching approaches). In the current era of information technology with a great amount of data available for nearly all facets of sports, automatic and rapid assessment systems for tactical behaviour in team sports are highly desirable. Use of (sports-) scientific methods has provided only very limited scope for investigation in the past. However, collaboration between sports sciences and computer science is expected to present considerable opportunities for synergistic benefits in the future. Future research should also try to focus on a detailed comparison of the three different type of outcomes resulting from the three different approaches. In further analyses, for example, our findings on tactical performance could be directly compared with the tactical performance of the defense or offensive players in some critical moments or phases. But these investigations would also need a general theoretical framework in performance analysis [47]. While this would be a challenging task it could foster performance analysis approaches based on position data from a theoretical and practical point of view.

Acknowledgements The content of this leading article is based on three lectures given by the authors to an invited symposium at the European College of Sport Science 2014 Congress held in Amsterdam, The Netherlands.

Compliance with Ethical Standards

Funding No financial support was received for the planning or conduct of the research presented in this article. Preparation of the article was supported by a Grant from the German Research Council (DFG, Deutsche Forschungsgemeinschaft) [ME 2678/3-3] to Daniel Memmert.

Conflict of interest Daniel Memmert, Koen A.P.M. Lemmink and Jaime Sampaio declare that they have no conflicts of interest relevant to the content of this article.

References


