

Article

Not peer-reviewed version

A Mathematical Model to Study Defensive Metrics in Football: Individual, Collective and Game Pressures

[Jose M. Calabuig](#), [César Catalán](#), [L.M. García-Raffi](#)^{*}, E.A. Sánchez-Pérez

Posted Date: 23 November 2024

doi: 10.20944/preprints202411.1787.v1

Keywords: Mathematical Modeling; Approximation Functions; Football metrics; Pressure model; Tracking data analysis



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A Mathematical Model to Study Defensive Metrics in Football: Individual, Collective and Game Pressures

J.M. Calabuig , C. Catalán , L.M. García-Raffi *  and E.A. Sánchez-Pérez 

Instituto Universitario de Matemática Pura y Aplicada. Universitat Politècnica de València. Camí de Vera s/n. 46022 València. Spain

* Correspondence: lmgarcia@mat.upv.es; Tel.: 0034-9638776600

Abstract: Performance analysis, utilizing video technology and recent technological advancements in soccer stadiums, provides a wealth of data, including player trajectories and real-time game statistics, which are crucial for tactical evaluation and decision-making by coaches and players. This data allows for the definition of metrics that not only enriches the experience for soccer fans through enhanced visual displays, but also empowers coaching staff and managers to make informed, real-time decisions that directly impact match outcomes. Ultimately, it serves as a pivotal tool for improving team strategy based on comprehensive post-match data analysis. In this article, we present a mathematical model to study the concept of pressure between players and, subsequently, between teams. We first explore the concept in a fixed frame of a match, determining what we call *influence areas* between players. We introduce the unit pressure function and analyze the total number of pressure interactions. Then, we apply these concepts to football matches, considering various factors such as players and the radius of the area of influence, examining pressure efficiency through mean unitary pressure. Lastly, a real case study is presented, showcasing visualizations like a heatmap matrix displaying individual and collective pressure, as well as team pressure balance.

Keywords: mathematical modeling; approximation functions; football metrics; pressure model; tracking data analysis

1. Introduction

Top-level soccer is an increasingly competitive sport and, as a consequence, performance analysis is becoming an important tool for supporting coaches and players in the decision-making process, providing relevant information about the play. For decades, the use of video has been a key element in the tactical analysis of teams before and after a match [1]). The technological advances developed in recent years in soccer stadiums now make it possible to obtain a large amount of data related to the game, both individually and collectively. These data, which contain the trajectories of the players, the ball or the movement of the refereeing team in real time, allow a very precise mathematical analysis of the parameters that characterize the course of the matches. In addition, broadcasters have realized that coverage of sporting events can be enhanced by providing analysis of the action, often by specialized professionals who use visual tools and dashboards to describe key events and sequences during the game. There is also an apparent consumer appetite for statistics and attractive visual data displays, and there are many websites dedicated to providing this type of information. Therefore, there is a general interest in improving the visualizations and filling them with technical content that can provide more information to soccer fans. But this is not only of interest to the fans, but also to the coaching staff and managers, who can analyze the data, even in real time, and make relevant decisions that will affect the outcome of the match. It is thus intended to give a tool for improving the strategy of the teams, based on real data, after processing the complete data describing the match.

The basic ways to obtain the real trajectories and positions of objects and players are mainly achieved by Electronic Performance and Tracking Systems (EPTS). Technologies based on Optical Tracking System (OTS) capture and can analyze 25 frames per second, which would already add up to a total of 135.000 sets of values for 90 minutes of play per game, and if we consider 22 players, a game ball and the referee team constitute a big-data of the order of $3 \cdot 10^6$ data vectors in \mathbb{R}^3 euclidean space

to be stored, which are not very feasible to analyze without an adequate modelling and treatment to obtain relevant and useful information for potential customers of use.

There are widely used indicators such as shots on goal, number of passes, tackle rates, team ball possession and distances covered although their use is a matter of question [2]. Then, it is a task of data science to derive indicators from raw data that describe relevant situations of the game properly ([3]). In particular, a lot of studies have analysed tactical indicators related to goal scoring in soccer (see for example [4]). In some of them, one of the concepts that recursively appears when we talk about football is the concept of pressure. Several attempts to define pressure can be found in the literature. In 1996 [5] introduce the *minimum moving time pattern (MMT)* of a player at any one moment as a pattern in which each point has the minimum time to move from that player's current position to that point; and also defined the *dominant region* of a player at any one moment as a region where the player can arrive at earlier than all of the others.

The concept of playing style is discussed in the reference [6], with special emphasis on what variables can be measured to determine it. In particular in attacking or offensive phases of play players move into positions towards the zones of the field in order to receive the ball or score a goal and in this phases of the play when pressure, considered as limiting the passing options and limiting space available, becomes fundamental.

One of the most characteristic and essential elements of soccer defense is pressing but it has not yet received much attention from researchers in football analytics. In [7] authors analyze the automatic extraction of football events using tracking data that implies among others fact the definition of the possession of the ball to determine in-game events. For this purpose they develop a decision tree-based algorithm and in the conclusions they establish as an improvement of their algorithm *extending the possession zone definition to encompass a variable radius/shape based on pitch location, proximity of opponents and player velocity*. The concept of pressure (which is directly related to that of possession) proposed in this article is based on the concept of this zone around the player and addresses, among other things, the determination of this radius.

[8] proposed a computational approach to detecting and quantifying the relationships of pressure and pressing behavior emerging during a game. To support examination of team tactics in different situations, they designed and implemented an interactive visual tool named *time mask* that enables selection of multiple disjoint time intervals in which given conditions are fulfilled.

In 2019 [9] studied as well an approach to quantification of attacking performance in football. Their procedure determined a quantitative representation of the probability of a goal being scored for every point in time at which a player is in possession of the ball they referred to this as *dangerousness*. This parametric variable was determined by means of four variables, one of them denoted as *pressure* which represents the possibility that the defending team prevent the player from completing an action with the ball. In determining the pressure, they assumed that a defender exerts pressure when his distance from the player is below some threshold value. They provided a pressure zone with covers four sub-areas with different radii which result of the angle between player and the centre of the goal. This model for defining *pressure* results in a rather complex formula.

In [10], in the context of tactical key performance indicators, authors introduced the Pressure Passing Efficiency Index (PPEI). It is based on the number of outplayed opponents. It aims to measure high quality through-balls by weighing passes with more than one outplayed opponent by the pressure on both pass initiator and receiver.

In [11] author introduces a novel metric to quantify its effectiveness in different contexts. He assess pressure calculating probabilities of recovering the ball and conceding a goal-scoring opportunity in the near future. They XGBoost stacked with a logistic regression model to estimate probabilities.

In [12] authors explore the area of defensive performance in football by analyzing successful defensive phases using space and time characteristics of defensive pressure. They use data from German Bundesliga games. The study distinguishes pressure on the ball-carrier, the group of attackers

closest to the ball, and the entire team, finding that pressure is higher in areas closer to the ball and as the defensive play approaches its end.

In this paper we have applied our definition of pressure to the analysis of data from a real football match. These data have been obtained from *Metrica company* ([13]). These are two games which have been anonymized, meaning there are no references to the names of players, teams or competitions. The dimensions of the field are the same for both games: 105×68 meters and tracking and event data are synchronized. A detailed information is given about the definition and explanation of all events types and subtypes.

The article is organized as follows. After this introductory section in the second one we first study the notions of pressure on a fixed frame of the match. To do that we start by introducing the concept of influence area (\mathcal{A}) between two players. This function depends of the distance between the players and also on the ball radius. An approximation of this function (by using a secant line) gives the definition of pressure (\mathcal{P}) between the players. This notion allows us to define the pressure between one team over a player and finally between the two teams. Since the cumulative pressure values can be very large then we introduce the unit pressure function (p) which it is computed by dividing the pressures by its maximum. Then we study the total number of pressure interactions. In the final part of this section we apply the previous formulas to study the pressure on a complete match of football. Now the pressure depends, not only on the players and on the ball radius, but also on the frame. We divide this section in three parts. In the first one we determine a ball radius in order to have a significant (but not too many) number of pressure interactions. Once this value ($r = 5\text{ m}$) is fixed we study the pressure along a set of frames (\mathcal{F}): the key point here is that we sum the pressures of all players defending the ball. Since in a match not all the players play the same time in the third part of this section we study the efficiency of the pressure by introducing the mean unitary pressure (\hat{p}). This provides a number in $[0, 1]$ easier to use. In the third section we present a real case study providing different visualizations using our formulas. Hence we start by showing a heatmap matrix that allows us to give (in the same table) the individual and the collective pressure (both exerted and received). In the second part a Butterfly Chart (also called Tornado Chart) is presented. This visualization allows us to study the balance of pressures of the teams of the match. In the last section we discuss the results and arrive to some conclusions.

2. Materials and Methods

2.1. Pressure Functions on a Frame of a Match

Let A and B the two teams on the pitch. In this section we first introduce the notion of pressure between the players $a \in A$ and $b \in B$ and also the corresponding one between one team on a player and also between the two teams. In order to simplify the computations an approximation of these functions—using a secant line—will be also given. Next we introduce the unitary concepts of pressure and we finish this section by providing the total number of pressing interactions.

Throughout this section we set a frame so there is no time dependence and the definitions depend only on the radius of the ball. In the next section we will use our results to analyze a complete match—with the corresponding time/frame dependence—and we will also look at how to determine the radius to compare the pressure between players and teams. We start by defining the notion of influence area between two players.

Definition 2. Fixed a real number $r > 0$. We define the **influence area between the players $a \in A$ and $b \in B$** , as the truncated area of the intersections balls $B(a, r) \cap B(b, r)$. That is,

$$\mathcal{A}_{a,b}^r = \mathcal{A}^r(d_{a,b}) := \begin{cases} 2r^2 \arccos\left(\frac{d_{a,b}}{2r}\right) - \frac{d_{a,b}}{2} \sqrt{4r^2 - d_{a,b}^2}, & 0 \leq d_{a,b} \leq r, \\ 0, & d_{a,b} > r, \end{cases} \quad (1)$$

where $d_{a,b}$ is the euclidean distance, $d_{a,b} = \|a - b\|_2$, and a and b stands for the coordinates of each player—in \mathbb{R}^2 .

Note that the objective of making the influence area function equal zero when $d_{a,b} > r$ is to take into account that when two players are far away enough this area cannot be considered. Figure 1 shows the truncated intersection areas between the balls $B(a, r)$ and $B(b, r)$ —that is the influence area between the players a and b —.

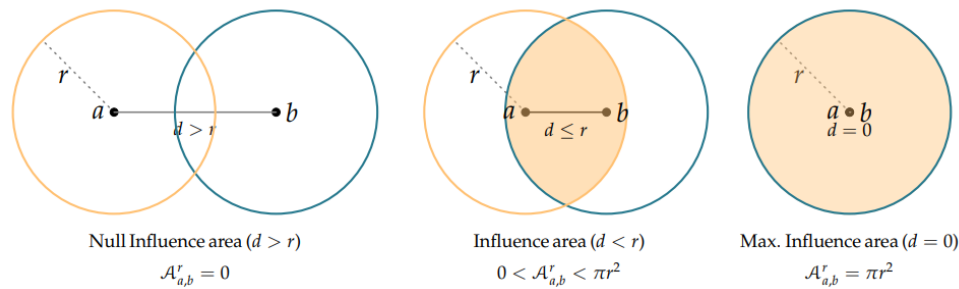


Figure 1. Influence area between players $a \in A$ and $b \in B$ depending on the radius r .

Thinking on the real-time computations we can approximate the influence area function $\mathcal{A}_{a,b}^r$ on the interval $[0, r]$ by the secant line from $(0, \mathcal{A}_{a,b}^r(0))$ to $(r, \mathcal{A}_{a,b}^r(r))$. This provides our notion of pressure.

Definition 2. We define, for $r > 0$, the *pressure exerted by player $a \in A$ over the player $b \in B$* as

$$\mathcal{P}_{a,b}^r = \mathcal{P}^r(d_{a,b}) := \begin{cases} \pi r^2 - \left(\frac{\pi}{3} + \frac{\sqrt{3}}{2}\right) r d_{a,b}, & 0 \leq d_{a,b} \leq r, \\ 0, & d_{a,b} > r. \end{cases} \quad (2)$$

The linear approximation provided in this case gives accurate results. In fact, it allows us to approximate the influence area function with small errors. Figure 2 shows a visualization of how the influence area and the pressure functions behaves for different values of r . With all this approximations we will speed up the calculations so that we can display the results dynamically in near real time.

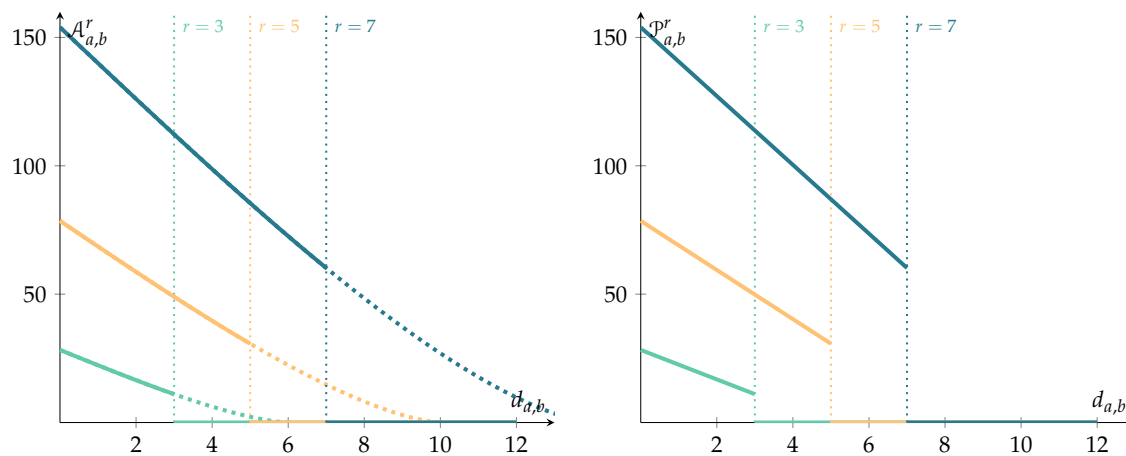


Figure 2. Representation of the influence area function—left—and its approximations—right—for different ball sizes $r=3, 5$ and 7 —as a function of the distance $d_{a,b}$.

In a similar way we can define the pressure between team and player and between two teams.

Definition 2. Given $r > 0$, the *pressure exerted by team A over the player $b \in B$* is defined as

$$\mathcal{P}_{A,b}^r = \sum_{a \in A} \mathcal{P}_{a,b}^r. \quad (3)$$

Definition 2. Fixed $r > 0$, the *pressure exerted by team A over the team B* is defined by

$$\mathcal{P}_{A,B}^r = \sum_{b \in B} \mathcal{P}_{A,b}^r = \sum_{b \in B} \sum_{a \in A} \mathcal{P}_{a,b}^r. \quad (4)$$

Note that some (actually many) players in $a \in A$ will exert zero pressure—over the player $b \in B$ —. In this case $\mathcal{P}_{a,b}^r = 0$, and then (many) terms of the sum will be null.

Sometimes, in order to show the results and make them more understandable—note that the cumulative values can be, in general, very large—, it is interesting to use the **unit pressure**, which is denoted by the small letter $p_{a,b}^r$. To define it recall the values of the pressure for the three singular situations given by $d_{a,b} = 0, r$ and $2r$ are

$$\begin{cases} \mathcal{P}_{a,b}^r(0) &= \pi r^2 =: f_{max}, \\ \mathcal{P}_{a,b}^r(r) &= \left(\frac{2\pi}{3} - \frac{\sqrt{3}}{2}\right)r^2 \approx 0.391 f_{max}, \\ \mathcal{P}_{a,b}^r(2r) &= 0 =: f_{min}. \end{cases} \quad (5)$$

To define the unit pressure it is enough to divide the pressure function by its maximum value, f_{max} —that is the area of the circumference of radius r —:

$$p_{a,b}^r = \frac{\mathcal{P}_{a,b}^r}{f_{max}} = \frac{1}{\pi r^2} \mathcal{P}_{a,b}^r. \quad (6)$$

We finish this section by counting the number of players of a team exerting pressure—in the sense that de pressure is different to zero—.

Definition 2. Let $b \in B$. We will write $n_{A,b}$ for the number of players on team A who exerts (non null) pressure over the player $b \in B$. In other words it is defined by:

$$n_{A,b} = \text{card}(\{a \in A : \mathcal{P}_{a,b}^r \neq 0\}). \quad (7)$$

We define the **total number of pressure interactions on the pitch** as

$$N = \sum_{b \in B} n_{A,b}. \quad (8)$$

Finally just observe that, the total number of interactions computed from each side—the pressure that all the members of B receive from the players of A , and the pressure that the members of A exert over the team B —coincide. This can be understood—and proved—by applying graph theory just thinking on two disjoint sets representing a graph consisting of $n_A + n_B$ vertices and N edges—see Figure 3—.

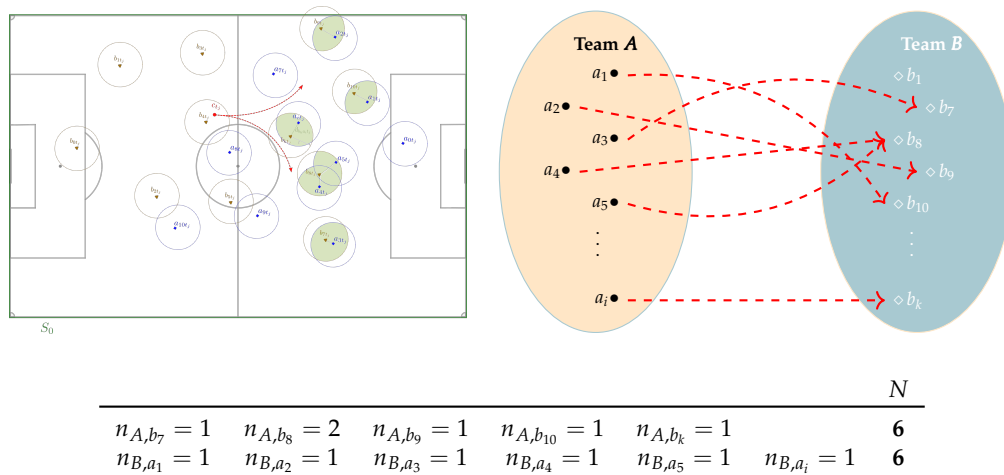


Figure 3. Example of total number of pressure interactions between the two teams on a frame.

2.2. Modelling the Pressure in a Complete Football Match

Now we will use the formulas introduced in the previous section in order to analyze the pressure in a complete match. As a starting point we take a set of frames \mathcal{F} . Note that now our pressure definitions depend on the frame $f \in \mathcal{F}$ —since the distance between a and b depends actually of $f \in \mathcal{F}$ —. To simplify our notation we write

$$\mathcal{P}_{a,b}^r(f) := \mathcal{P}^r(d_{a,b}(f)). \quad (9)$$

Of course if any of the players are not on the pitch in the frame $f \in \mathcal{F}$ then $\mathcal{P}_{a,b}^r(f)$ is defined as zero.

STEP 1. Determining the radius. In order to obtain a useful tool to analyze a football match, first we want to determine which values of r could be optimal to obtain a metric. Let us take the first part of the game which is composed of a total of 70.578 frames. The origin of these data was commented in Section 1. Figure 4 shows the positions on the pitch—in a particular frame $f \in \mathcal{F}$ —for four different radii— $r = 1, r = 4, r = 7$ and $r = 10$ —and the corresponding total number of pressure interactions.

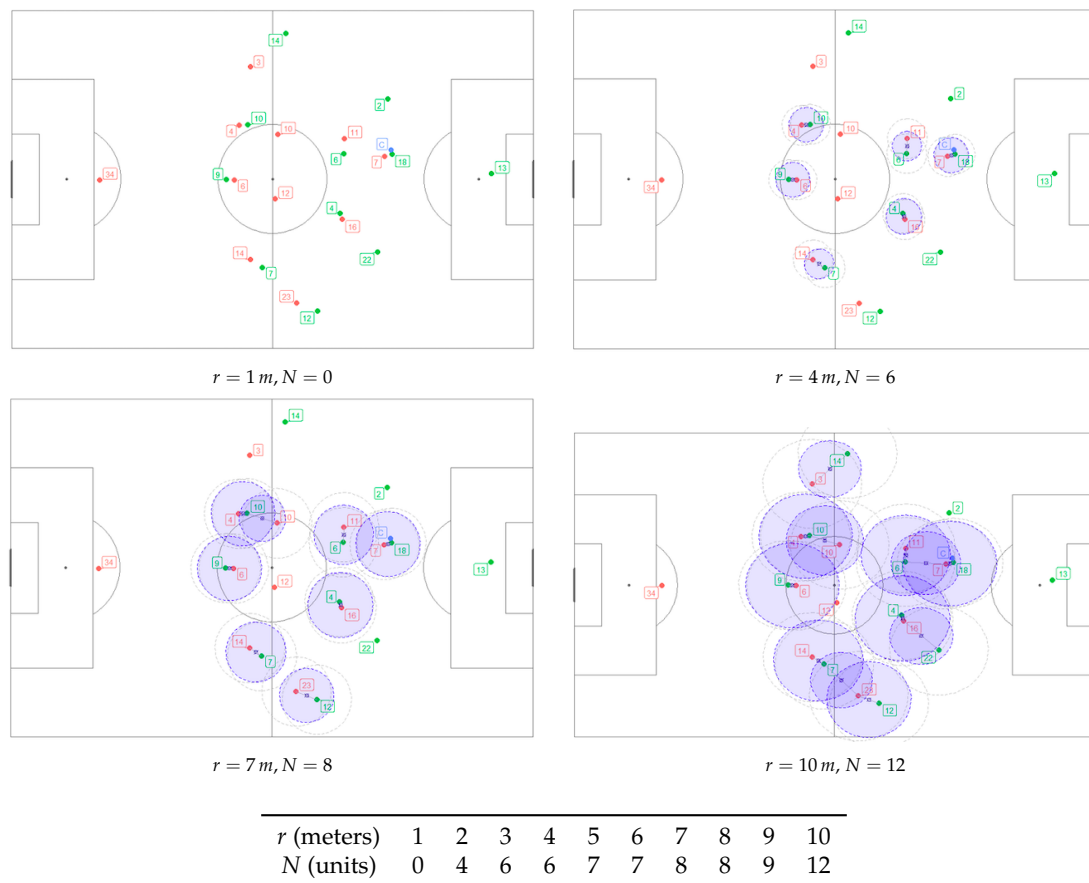
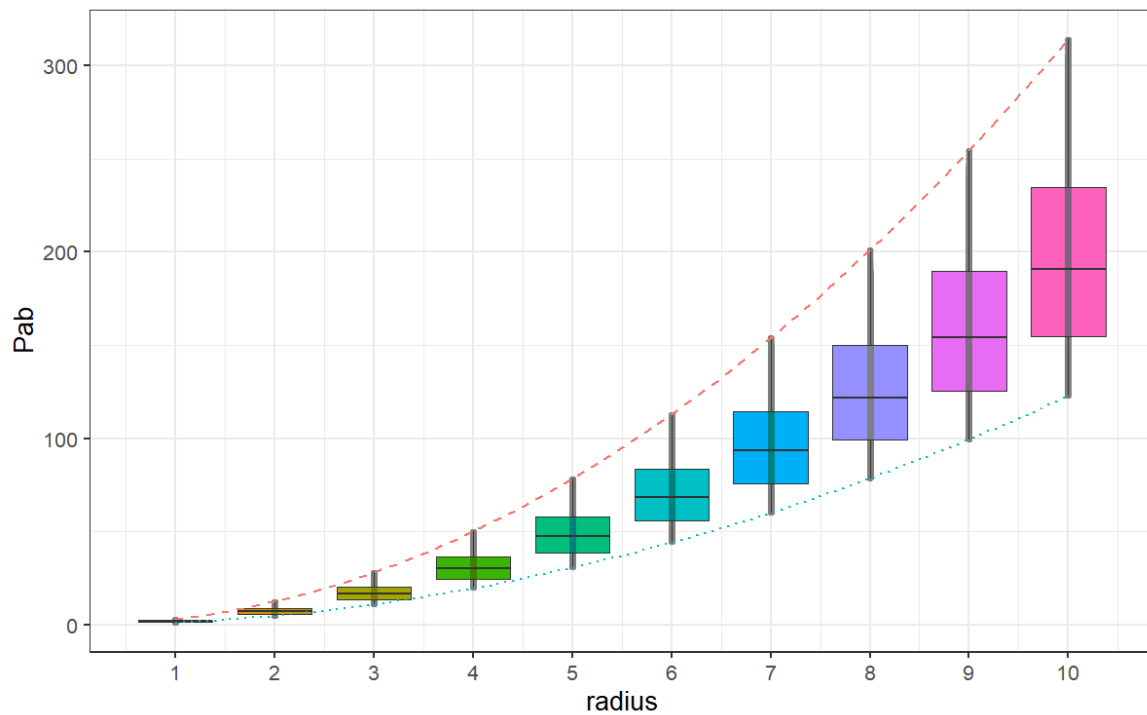


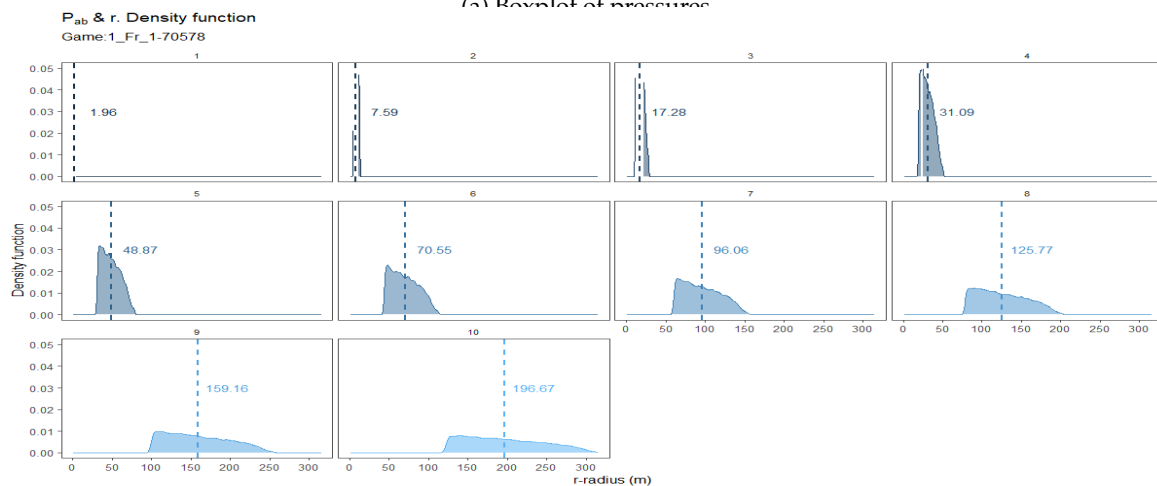
Figure 4. $\mathcal{P}_{a,b}^r$ sensibility to the ball radius variation in a frame $f \in \mathcal{F}$.

In the general case we want a ball radius $r > 0$, which provides a large number of interactions $N(\mathcal{F})$ —of course depending of the set of frames \mathcal{F} —, but avoiding that the associated pressures have large oscillations. The decision on what is the optimal value of the ball radius have to be the result of a compromise between: (a) a significant number of interactions N to allow accurate estimates, and (b) avoid considering distant players as participants in the pressure. The pressure depends on the number of interactions, as explained above. The data obtained shows quite a few outliers, with higher values: the number of high outliers is due to the frames belonging to corners and throw-ins where players tend to group together more.

Figure 5a shows a Box and Whisker plot for the pressures depending on the value of r and also on the frame set \mathcal{F} . The dotted lines show the $\mathcal{P}_{a,b}^r(0) = f_{max}$ of the pressure and the $\mathcal{P}_{a,b}^r(r) = 0.391f_{max}$ value. It can be seen that the inter quartile range increases considerably as the radius r increases. Moreover, if we look at Figure 5b, where we plot the density functions obtained by the sampling of the corresponding values of $\mathcal{P}_{a,b}^r$, we observe that the distributions flatten considerably as r increases. In principle, excessive flattening is not in our interest to obtain a good model, since too much dispersion would lead to a lack of precision in the results.



(a) Boxplot of pressures.



(b) Density of pressures.

Figure 5. Pressure distributions for different radii along the set of frames \mathcal{F} .

In order to reach a balance between high values of N —now along the set of frames \mathcal{F} —and concentrated distributions for the pressure we chose to use the shape factors of the distributions, namely the Curtosis coefficient. In Figure 6 we show the plot of the normalized mean of the interactions for each value of r and the normalized Curtosis values for each of these radii. As we see we reach the desired equilibrium point for the model for a radius on the interval $[4, 6]$ —we can take, for instance, the middle point $r = 5 m$ —.

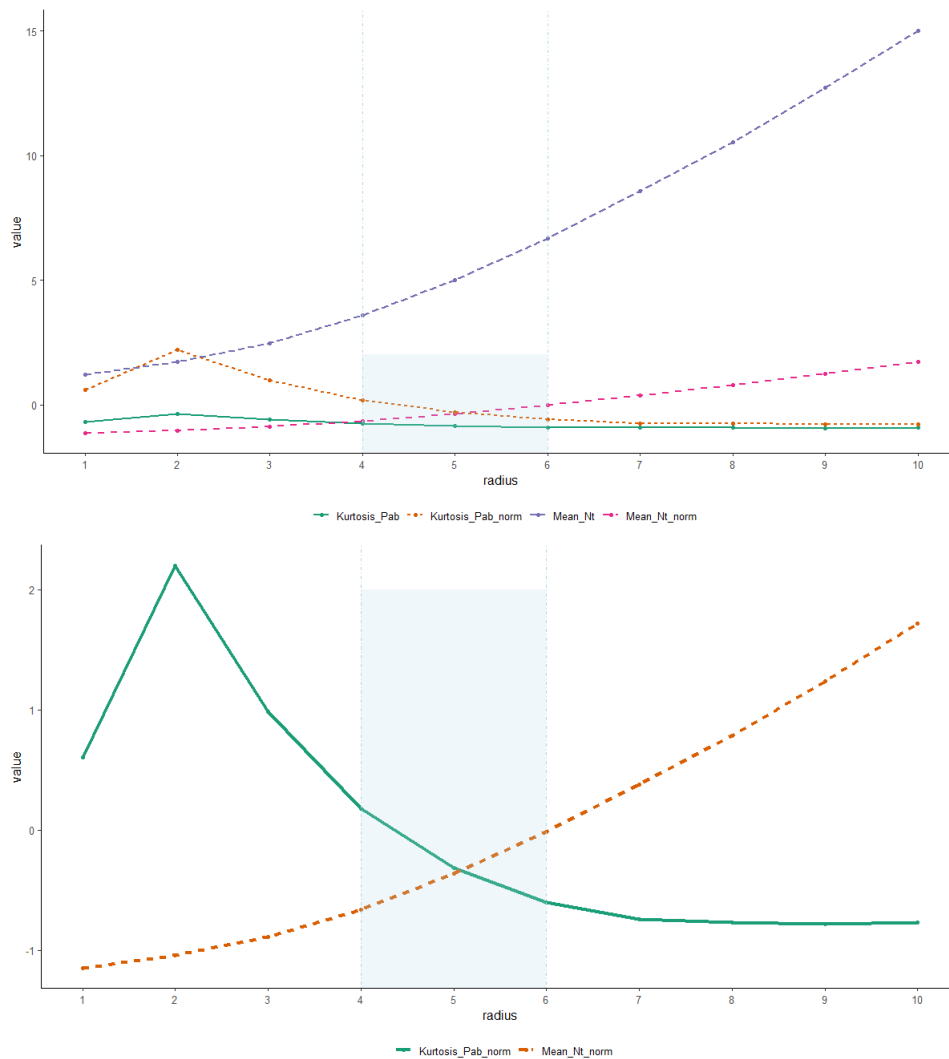


Figure 6. Optimal ball radius using the normalized mean of the interactions and the normalized Kurtosis of the number of interactions between players. We set the radius as $r = 5$ meters.

STEP 2. Studying the pressure. Once the parameters of the model and the corresponding approximation has been done, we want to use our pressure formulas in order to study the pressure between players and teams on the complete match. First we fix a set of frames \mathcal{F} . **Let us assume A is the defending team**—and therefore B has possession of the ball—. Then we write that—the players of— A exert pressure over—the players of— B , or equivalently—the players of— B receive pressure over—the players of— A . Hence we make a disjoint partition $\mathcal{F} = \mathcal{F}_A \cup \mathcal{F}_B$ where \mathcal{F}_A are the frames where A defends—i.e. does not have the possession of the ball—and \mathcal{F}_B the corresponding ones where team B is defending.

Definition 2. Let $r > 0$ and \mathcal{F} a set of frames. We define the **pressure exerted by the player $a \in A$ over the player $b \in B$ along the set \mathcal{F}** as

$$\mathcal{P}_{a,b}^r(\mathcal{F}) = \sum_{f \in \mathcal{F}_A} \mathcal{P}_{a,b}^r(f),$$

where $\mathcal{P}_{a,b}^r(f)$ stands for the pressure on the frame f —see (9)—.

Example 2. Suppose we take a set $\mathcal{F} = \{f_1, f_2, f_3\}$ where in f_1 and f_2 the team A defends—so B has the possession of the ball—and in f_3 is the team B which defends—see Figure 7—.

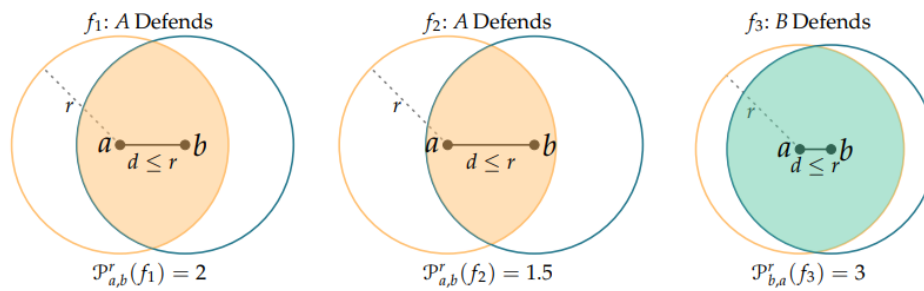


Figure 7. Example of pressure with a set of three frames.

Then we have that

$$\mathcal{P}_{a,b}^r(\mathcal{F}) = \mathcal{P}_{a,b}^r(f_1) + \mathcal{P}_{a,b}^r(f_2) = 3.5 \quad \text{and} \quad \mathcal{P}_{b,a}^r(\mathcal{F}) = \mathcal{P}_{b,a}^r(f_3) = 3.$$

In a similar way we define the **pressure exerted by the team A over the player $b \in B$ along the set \mathcal{F}** as

$$\mathcal{P}_{A,b}^r(\mathcal{F}) = \sum_{a \in A} \mathcal{P}_{a,b}^r(\mathcal{F}) = \sum_{a \in A} \sum_{f \in \mathcal{F}_A} \mathcal{P}_{a,b}^r(f).$$

Note that $\mathcal{P}_{A,b}^r(\mathcal{F})$ can be also seen as the **pressure received by the $b \in B$ along the set \mathcal{F}** . Also the **pressure exerted by the team A over the team B along the set \mathcal{F}** (or the received by the team B along the set \mathcal{F}) is given by

$$\mathcal{P}_{A,B}^r(\mathcal{F}) = \sum_{b \in B} \mathcal{P}_{A,b}^r(\mathcal{F}) = \sum_{b \in B} \sum_{a \in A} \mathcal{P}_{a,b}^r(\mathcal{F}) = \sum_{b \in B} \sum_{a \in A} \sum_{f \in \mathcal{F}_A} \mathcal{P}_{a,b}^r(f).$$

STEP 3. Efficiency of the pressure. Our starting point here is that not all the players play the same time—there are players that are sent off or substituted—. To deal with this problem we study the **mean unitary pressure** notion. In a sense this try to explain the effectiveness of a pressure. Since the cumulative values of the individual exerted pressures during the whole match can be very high, one can use instead the unit pressures, $p_{a,b}^r$, as defined in (6). More precisely:

$$p_{a,b}^r(\mathcal{F}) = \frac{\mathcal{P}_{a,b}^r(\mathcal{F})}{\pi r^2} = \frac{1}{\pi r^2} \sum_{f \in \mathcal{F}_A} \mathcal{P}_{a,b}^r(f). \quad (10)$$

Recall that we are assuming that A is the defending team—so the ball possession corresponds to B —, and \mathcal{F}_A are the set of frames where A defends. Hence, let us write \mathcal{F}_A^a , to the set of frames where A is defending and player $a \in A$ is on the pitch. Note that the cardinal of this set corresponds to the time where a is on the pitch and team A defends. Actually since in our database each *second* has 25 frames then the time is $t_A^a = 25 \text{card}(\mathcal{F}_A^a)$. Therefore:

Definition 2. Let $r > 0$ and \mathcal{F} a set of frames. We define the **mean unitary pressure exerted by the player $a \in A$ over the player $b \in B$ along the set \mathcal{F}** as

$$\hat{p}_{a,b}^r(\mathcal{F}) = \frac{p_{a,b}^r(\mathcal{F})}{\text{card}(\mathcal{F}_A^a)} = \frac{1}{\text{card}(\mathcal{F}_A^a)} \frac{1}{\pi r^2} \sum_{f \in \mathcal{F}_A} \mathcal{P}_{a,b}^r(f). \quad (11)$$

Note that $\hat{p}_{a,b}^r$ is a number on the interval $[0, 1]$.

3. Results: A Real Case Study: Visualizations

In this section we study the pressure exerted by the teams in one entire match of the anonymized game (Sample Game 1) offered by ([13]). We propose different visualizations for representing our notions. Visualization is of primary importance to make our results useful. As we said before since the cumulative pressures are very large we prefer to use the corresponding unitary pressure values—that is the functions $p_{a,b}^r$.

The first visualization proposed is based on the construction of heatmap-matrices. In Figure 8 we present the two matrices: in the figure above team A defends and in the figure below is team B the defender team. The interpretation of the tables are as follows:

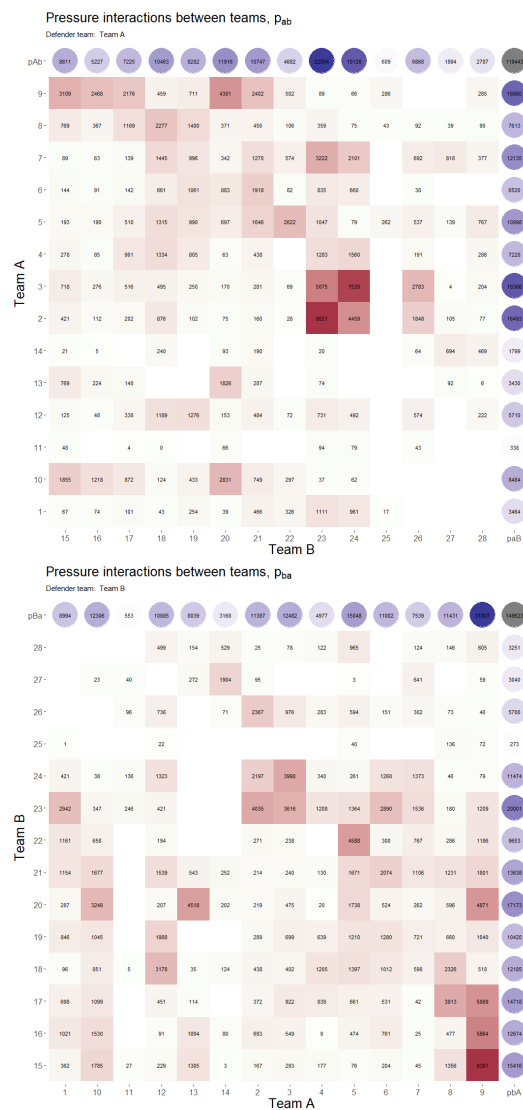


Figure 8. Individual and collective pressure unitary metrics. Above: Exerted by team A and received by team B. Below: Exerted by team B and received by team B.

- Individual exerted unitary pressure.** The i row of the matrix gives the individual pressure exerted by the corresponding labelled player—marked in red colors—. For instance, the three players who have exerted the most pressure when team A defends—figure above—are the ones labelled with the numbers 2 (8.021 units of unitary pressure against the player 23 and 4.459 units

of unitary pressure against the player 24), the player 3 (7.520 units of unitary pressure against the player 24 and 5.075 units of unitary pressure against the player 23) and the player 9 (4.301 units of unitary pressure against the player 20).

2. **Collective exerted unitary pressure.** The last column—blue balls—explains the collective exerted pressure of each player. In this case the three players exerting the most collective pressure—also for the figure above—are the numbers 3 (18.366 units), 9 (16.860 units) and 2 (16.493 units).
3. **Collective received unitary pressure.** In the first row—also blue balls—we can see the total pressure received for each player—of the attacking team—. Again in the figure above the player receiving the most collective pressure are the numbers 23 (22.004 units), 24 (18.120 units) and 20 (11.916 units).
4. **Unitary pressure exerted along the match.** Finally a grey ball—on the right corner of the table—gives de total exerting pressure of the match—which coincides with the total received pressure—. When team A defends is a total amount of 119.443 units—figure above—being 149.623 units when is B the defender team—figure below—. Note that the unit pressure has been calculated with a precision of 4 decimal places, so there may be small discrepancies in the total sum of the unit pressures with respect to the teams, for example, when team A is defending the respective sums of the top row and right column give values of 119.435 and 119.438 respectively when the sum made directly through the unit pressures amounts to 119.443 (grey box), reaching an maximum absolute non-significant error of 8 units ($6,70 \cdot 10^{-5}$)

Recall that a value of 1.000 units in the unit pressure would be equivalent to a player having been in total contact with another player along 1.000 frames, i.e., 40 seconds.

Regarding the (unitary) collective pressure—both exerted and received—we can also present a new metric: the defensive balance. Again we are using the unitary pressure, $p_{a,b}^r(\mathcal{F})$, as defined in (10). We have chosen a Butterfly Chart (<https://datavizproject.com/data-type/butterfly-chart/>) as it is shown in Figure 9. In order to do that, we assign to each player a minus sign to the total received unitary pressures—representing it with red color—and a plus sign to the corresponding total exerted unitary pressure—now with blue color—. Therefore the **defensive balance** is given by the difference between the exerted and the received unitary pressure. In Figure 9 it can be seen that A team has a global negative balance of pressure in the match: a total amount of $-30\,180$ units—see the blue box of the figure above which is the difference between the exerted unitary pressure (119.443 units) and the received unitary pressure (149.623 units)—.

But, much more information can be obtained from this representation. In the figure above we can see from each player of A team the total exerted unitary pressure—right part in blue color—and also the total received unitary pressure—left part in red color—. Analogously we see the same results from each player of B team—figure below—. Of course the exerted unitary pressure of a player (which is in the defender team) is the sum of the exerted unitary pressure between this player and all the players from the opposite team (which is suppose to have the ball). In a similar way we have the received unitary pressure. For example, we can see that the most pressured players where player 9 from A team (-31.807 units from figure above) and 23 from B team (-22.004 units from figure below). On the other hand player 3 from A team and player 23 from B team where the ones who exerted more pressure along the match (18.366 and 20.001 units of pressure respectively).

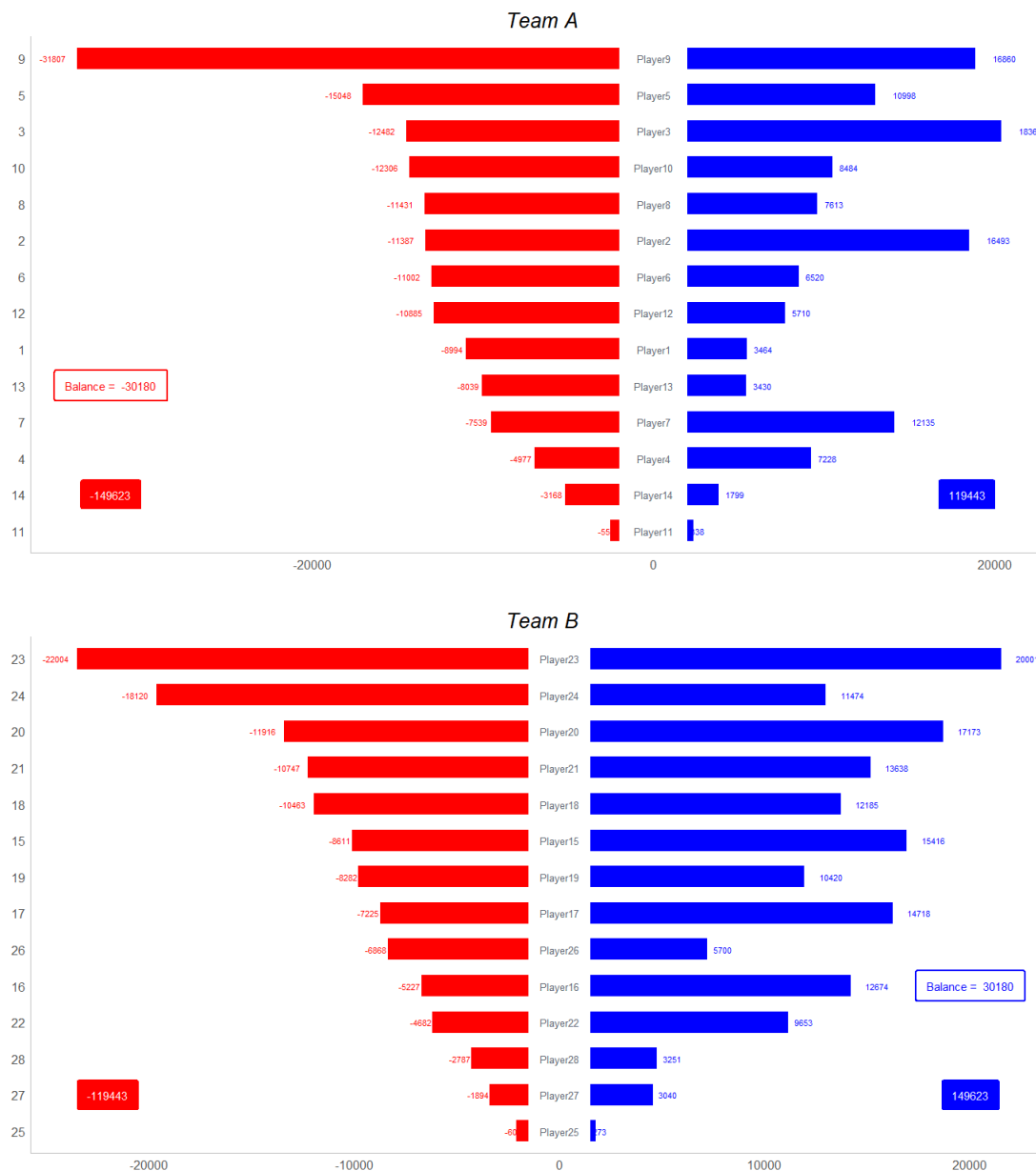


Figure 9. Defensive balance obtained from the unitary pressure $p_{a,b}^r(\mathcal{F})$. Above: Team A players against B Team. Below: Team B players against A Team.

We finish with the efficiency of the pressure. As we said before not all the players play the same time. Hence we use the mean unitary pressure notion obtaining the **mean defensive balance**. This is shown in Figure 10. As we explain before here we use the mean unitary pressure, $\hat{p}_{a,b}^r(\mathcal{F})$, as defined in (11) instead of $p_{a,b}^r(\mathcal{F})$.

During the match there were the following replacements:

- Team A: (1) \leftarrow (12), (6) \leftarrow (13), (10) \leftarrow (14).
- Team B: (22),(24) \leftarrow (26),(27), (19) \leftarrow (28).

Therefore comparing Figures 9 and 10 we see that, for instance, player 3 from A team is now the player with more exerted pressure (with value 0.4497), but player 9 (-0.6709) is not longer the most defended player but this position has passed to player 1 (-0.673). With respect to team B, player 23 is still the player with the best pressure (0.4219) but now closely followed by player 26 with value 0.3507. This means that player 26 of B team has done a great deal of pressing during the time he has been on the pitch.

In addition, through Figures 9 and 10 we can evaluate the goodness of the substitutions made in the match from the defensive point of view, that is, we observe that the unit pressures generated by player 12 (0.2100) are lower than those of player 1 (0.2537) which he has substituted but that of player 13 (0.3410) has been much higher than that of player 6 (0.2118) and that of player 14 (0.2519) somewhat higher than those of player 10 (0.2517) so that especially the substitution (6) \leftarrow (13) has been very productive for team A in terms of pressure generation.

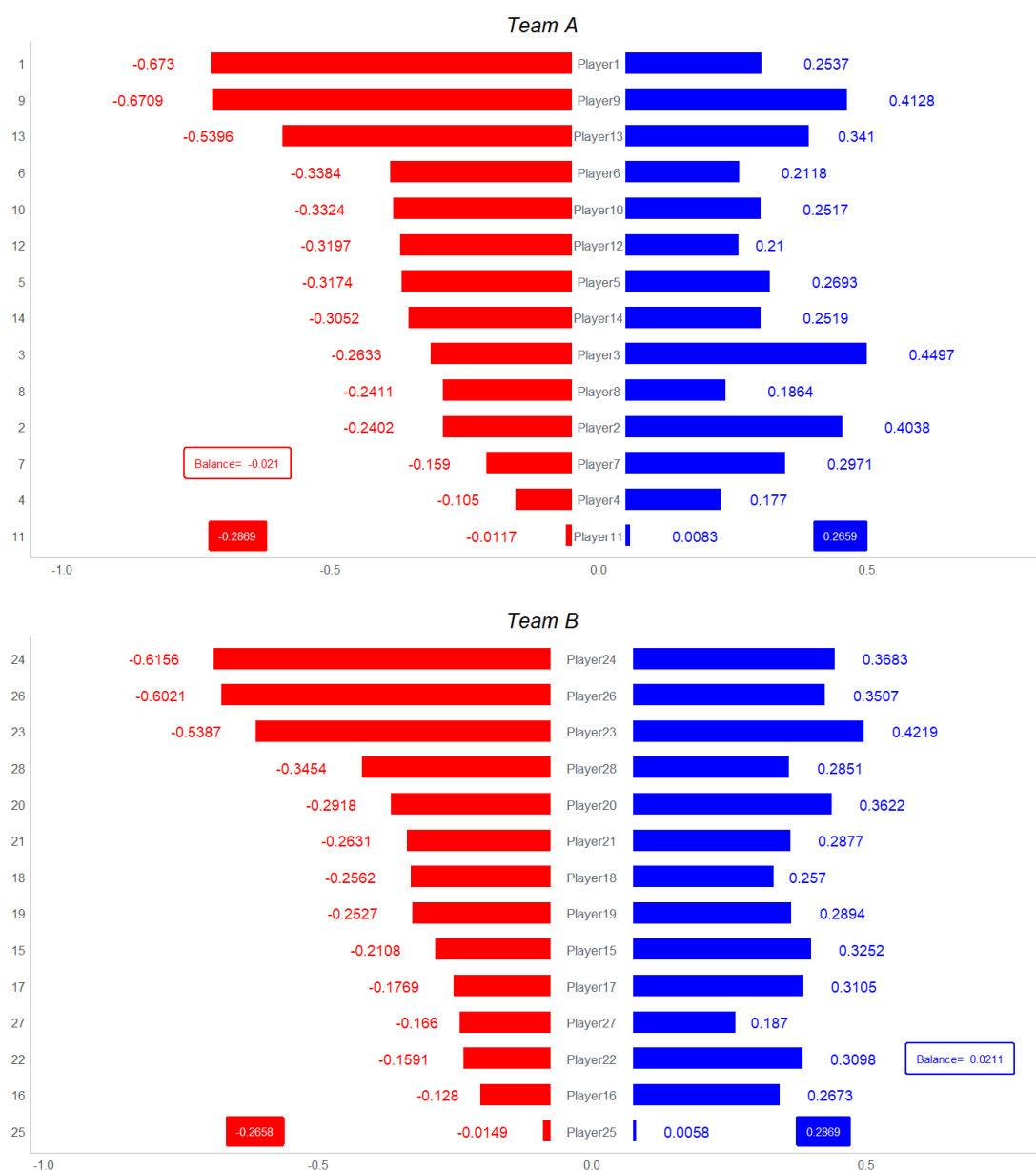


Figure 10. Mean defensive balance obtained from the mean unitary pressure $\hat{p}_{a,b}^r(\mathcal{F})$. Above: team A players against B team. Below: team B players against A team.

4. Discussion

In football, the pressure exerted both between players and between teams is a very important tactical performance indicator and it comes fundamental to design game strategies, determining changes in players' positions in the pitch, substitutions of players, spaces between lines...

Pressure evaluation is difficult because it is a metric that involves both individual players and groups of players. Therefore, having a simple and easily computable method to evaluate it from the data available in football can be useful for coaches and technical staffs.

In this paper we deal with a geometrical model of the definition of pressure. It starts from the elementary notion of area of influence, the area that is under the control of a player, which allows us to define the pressure exerted on a player based on the intersection of the areas of influence between the player who pressures and the player who is pressured. This distinction between the player exerting pressure and the player under pressure is made on the basis of which team is in possession of the ball at any given moment. This area of influence is established on the basis of a circumference centered on the player's position with radius r . This radius, a unique value for the complete match, is determined via a compromise between in one hand a significant number of interactions N to allow accurate estimates and on the other hand avoiding to consider very distant players as participants in the pressure. We want to note that this is the only parameter of the model whose determination from the data is straightforward as explained in the Subsection 2.2.

An important advantage of our model is that, once the elementary amount of pressure between players is defined, with the same model and the same assumptions, it is very easy to obtain simultaneously pressure received by a player from the whole opponent team and also pressure exerted by a team on the whole opponent team, everything using the same dataset information on individual players. In particular, we can evaluate the goodness of the substitutions made in the match from the defensive perspective analysing the pressure exerted by the player and by the substitute. It is that we call the efficiency of the pressure.

Our pressure metrics (individual and collective) are defined frame by frame. Therefore they can also be extended to the whole duration of a football match, which implies that not only can give an static game perspective but also a time-dependent perspective on a given match. As future work we plan to explore this time dependence via this detailed frame-by-frame analysis.

We have performed an analysis of data from a real match showing the versatility of this metric. For doing that we have used tracking data with eventing that allows us to determine at each moment the position of each player in the pitch, i.e. their relative positions, as well as to know which team is in possession of the ball at each moment. In particular we have shown that in a simple way a global balance can be obtained between two teams over the course of a complete match. This analysis can be restricted to any time interval within the match, determining the balance of pressures at different instants of the game. The same analysis can be done on a player-by-player on the same conceptual basis.

In summary, the model presented here is mathematically simple, easily computable and flexible in that it supports player and team, static and dynamic time-dependent metrics and we believe it can be a useful tool for the tactical analysis of football matches by coaches and technical teams, as well as broadcasters and audiovisual media.

References

1. O'Donoghue, P. The use of feedback videos in sport. *International Journal of Performance Analysis in Sport* **2006**, *6*, 1–14. doi:10.1080/24748668.2006.11868368.
2. Mackenzie, R.; Cushion, C. Performance analysis in football: A critical review and implications for future research. *Journal of Sports Sciences* **2013**, *31*, 639–676. doi:10.1080/02640414.2012.746720.

3. Arnau Notari, A.R.; Calabuig, J.M.; Catalan, C.; Garcia-Raffi, L.M.; Pardo Gila, J.M.; Pons Anaya, R.; Sánchez Pérez, E.A. Using neural networks and hierarchical cluster analysis to study goal kicks in football. *International Journal of Sports Science & Coaching* **2023**, *19*, 17479541231207184, [<https://doi.org/10.1177/17479541231207184>]. doi:10.1177/17479541231207184.
4. Smith, R.A.; Lyons, K. A strategic analysis of goals scored in open play in four fifa world cup football championships between 2002 and 2014. *International Journal of Sports Science & Coaching* **2017**, *12*, 398–403. doi:10.1177/1747954117710516.
5. Taki, T.; Hasewaga, T.; Fukurama, T. Development of motion analysis system for quantitative evaluation of teamwork in soccer games. Proceedings of 3rd IEEE International Conference on Image Processing, 1996. doi:10.1109/ICIP.1996.560865.
6. Hewitt, A.; Greenham, G.; Norton, K. Game style in soccer: what is it and can we quantify it? *International Journal of Performance Analysis in Sport* **2016**, *16*, 355–372. doi:10.1080/24748668.2016.11868892.
7. Vidal-Codina, F.; Evans, N.; El Fakir, B.; Billingham, J. Automatic event detection in football using tracking data. *Sports Engineering* **2022**, *25*, 18. doi:10.1007/s12283-022-00381-6.
8. Andrienko, G.; Andrienko, N.; Budziak, G.; Dykes, J.; Fuchs, G.; von Landesberger, T.; Weber, H. Visual analysis of pressure in football. *Data Min Knowl Disc* **2017**, *31*, 1793–1839. doi:10.1007/s10618-017-0513-2.
9. Link, D.; Lang, S.; Seidenschwarz, P. Real time quantification of dangerousity in football using spatiotemporal tracking data. *International Symposium on Computer Science in Sport* **2019**, pp. 1–16. doi:10.1371/journal.pone.0168768.
10. Memmert, D.; Raabe, D.; Schwab, S.; Rein, R. A tactical comparison of the 4-2-3-1 and 3-5-2 formation in soccer: A theory-oriented, experimental approach based on positional data in an 11 vs. 11 game set-up. *PlosONE* **2020**. doi:10.1371/journal.pone.0210191.
11. Robberechts, P. Valuing the art of pressing. <https://people.cs.kuleuven.be/~pieter.robberchts/repo/robberchts-statsbomb19-pressing.pdf>, accessed on 2024.
12. Forcher, L.; Forcher, L.; Altmann, S.; Jekauc, D.; Kempe, M. The keys of pressing to gain the ball – characteristics of defensive pressure in elite soccer using tracking data. *Science and Medicine in Football* **2022**, pp. 1–9. doi:10.1080/24733938.2022.2158213.
13. Metrica sports sample data. <https://github.com/metrica-sports/sample-data>, accessed on 2024.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.