



# OPEN The interplay between pitch control and top speed in soccer considering the stamina factor

Marco F. da Silva<sup>1,2</sup>, Álvaro Novillo<sup>3</sup>, Alberto Aleta<sup>1,2</sup>, Roberto López del Campo<sup>4</sup>, Ricardo Resta<sup>4</sup>, Yamir Moreno<sup>1,2</sup> & Javier M. Buldú<sup>3,5</sup>✉

In this study, we investigate the interplay between player speed and ball control in soccer. We present a novel pitch control algorithm that quantifies the probability of a player gaining possession at any location on the field. Our model accounts for the heterogeneity of player speeds by measuring performance during matches and assigning each player a specific top speed. We then compare the pitch control percentages derived from our approach with those from classical models, which assume uniform top speeds for all players, and analyze the results across different player roles (defenders, midfielders, and forwards). Our findings reveal a positive correlation between a player's top speed and their accumulated pitch control, with certain players benefiting more from this relationship. However, this positive correlation is constrained by the role of the player in the team, with defenders achieving the highest accumulated pitch control despite not being the fastest. Furthermore, our methodology supports team-level analysis, identifying which teams gain the greatest advantage from their players' top speeds, and extends to comparisons between the first and second halves of matches. Our model also enables exploration of how changes in top speed may affect pitch control at both the individual and team levels. To facilitate this, we introduce the stamina factor, a parameter that adjusts a player's top speed. We find that the impact of the stamina factor on pitch control follows a logarithmic function, with the scaling factor quantifying the potential benefits of increased speed. Interestingly, the influence of the stamina factor varies significantly by player position. Overall, our approach provides valuable insights into which teams or players could benefit most from improvements in physical performance.

**Keywords** Motion Models, Human Movement, Sports Analytics, Pitch Control

One of the applications of complexity sciences is the study of patterns movements of animals, herds, or humans, aiming to understand the underlying principles governing collective behaviors across species<sup>1–3</sup>. This interdisciplinary task has evolved over decades, with significant contributions that have deepened our understanding of how collective movement arises from the particular dynamics of its agents. In 1971, W.D. Hamilton introduced the “selfish herd theory,” indicating that individuals within a group position themselves to minimize their predation risk, leading to the emergence of cohesive group formations<sup>4</sup>. This foundational concept highlighted how individual actions could result in complex group dynamics. Building upon this, in 1995, Vicsek et al. developed a model demonstrating how simple behavioral rules could lead to the spontaneous emergence of collective motion in groups of self-propelled particles<sup>5</sup>. This model described the significance of local interactions in the formation of large-scale patterns observed in animal groups. Since this seminal work, a diversity of studies, covering a variety of disciplines, focused on identifying the laws describing the essential aspects of collective motion, discussing experiments, mathematical methods, and models for simulations<sup>3</sup>. Interestingly, all this literature points to the fact that, despite the multidisciplinary nature of the field, similar laws describe systems of completely different nature, ranging from macromolecules to groups of animals and people.

In recent years, the principles derived from the study of collective animal behavior have been increasingly applied to the analysis of human activities, particularly in sports. Team sports, such as soccer, exhibit complex dynamics where individual and collective movements are crucial for performance<sup>6</sup>. Understanding these

<sup>1</sup>Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Zaragoza, Spain.

<sup>2</sup>Department of Theoretical Physics, Faculty of Sciences, University of Zaragoza, Zaragoza, Spain. <sup>3</sup>Complex Systems Group, Universidad Rey Juan Carlos (URJC), Madrid, Spain. <sup>4</sup>Mediacoach, LaLiga, Madrid, Spain. <sup>5</sup>Grupo Interdisciplinar de Sistemas Complejos (GISC), Madrid, Spain. ✉email: javier.buldu@urjc.es

dynamics through the lens of complexity science has led to new perspectives on performance analysis and strategy development<sup>7,8</sup>. This approach acknowledges the intricate interactions and emergent behaviors that arise from individual actions within the team context<sup>9</sup>. Advancements in tracking technologies have enabled access to detailed data on player/athlete movements. Tracking datasets contain the precise Euclidean location of players during the different phases of training or competition, which, in turn, can be used to calculate their speed and acceleration. Sports such as basketball<sup>10,11</sup>, rugby<sup>12</sup> or soccer<sup>13</sup>, have benefited from the analysis of tracking datasets, extracting information about not only the physical performance of athletes and players but also the tactical implications of the individual and collective patterns of movement<sup>14,15</sup>.

Furthermore, tracking datasets have facilitated analyses that go beyond traditional statistics, allowing for a deeper understanding of spatial and temporal patterns in player behavior. For instance, in<sup>16</sup>, the location of professional soccer players during the attacking and defensive phases of the match was used to extract their corresponding speeds. The analysis of the interplay between player speed, movement angles, and distance to the ball, offered a new (and more complex) perspective on how players move strategically during different phases of the game. It was shown how the player speed is highly dependent on ball proximity, game phase (attack, defense), or player role, and how the angles of movement reveal role-specific behaviors that align with tactical responsibilities<sup>16</sup>. In a more general framework, the application of complexity sciences in soccer highlighted the intricate balance between individual decision-making and emergent group behaviors, reinforcing the idea that player movement is not just a sum of individual actions but rather an interactive, dynamic process influenced by teammates, opponents, and contextual factors such as game situations and tactical instructions. For example, in<sup>17</sup>, Marcelino et al. analyzed the correlation of the players' speeds to show that team coordination during soccer matches is not a fixed, pre-determined formation but rather an emergent, self-organized process. By analyzing players' spatiotemporal trajectories, they demonstrated that players continuously adjust their positions in response to both teammates and opponents, resulting in dynamically shifting clusters. These collective movement patterns change in response to different game phases, particularly between offensive and defensive moments<sup>17</sup>. On the other hand, in 2020, network science was used to analyze tracking datasets, and the concept of *tracking networks* was first introduced<sup>18</sup>. Under this framework, tracking datasets are used to quantify different kinds of interactions between players, leading to marking networks, coordination networks, or signed proximity networks. The organization of these complex networks along a match can be tracked and different network parameters can be computed to better understand the collective movements of soccer teams and players<sup>18,19</sup>.

In this paper, we focus on a specific application of the tracking datasets: the estimation of the pitch control of a soccer team. The idea of quantifying the pitch control was first introduced by Taki and Hasegawa<sup>20</sup> with the concept of *dominant regions*. Departing from the player location at every moment of a match, Taki and Hasegawa developed a mechanical model to simulate the movement of players and calculate the probability of controlling the ball at any future location of the pitch. In their model, the speeds and accelerations of players were included, going one step beyond the classical analysis using Voronoi partitions (i.e., disregarding the player dynamics) as a technique to evaluate the amount of the pitch controlled by each team<sup>21,22</sup>. Therefore, not only the proximity to the ball but also the direction of the speed is crucial to determine what player will arrive before to any location of the pitch to control the ball. Some year later, Spearman departed from the concept of *dominant regions* to develop a tracking-based model of passing performance<sup>23</sup>. Players' speeds, acceleration, reaction times, and positions on the field were integrated to obtain a Pass Success Probability Function. Combining the ball's flight characteristics with the players' motion capabilities, the Spearman model showed that the probability of a pass being successful is directly related to the "window" during which the intended teammate can receive the ball before an opponent can intercept. Furthermore, the predicted pass probabilities aligned well with the observed success rates from real match data.

One significant limitation of the models presented in<sup>20</sup> and<sup>23</sup> is the lack of heterogeneity in player speeds. Although these numerical models incorporated actual speed data from tracking systems, they assumed a uniform top speed for all players, despite evidence that top speeds vary among individuals<sup>24</sup>. While such a simplification may be justifiable in classical pitch control models, its effect on the true dynamics of pitch control for players and teams remains unclear. To address this gap, we have implemented a pitch control model that accounts for the actual top speeds of the players under evaluation. Using tracking datasets from 100 professional soccer matches, we first calculate each player top speed for every match and then integrate these values into our numerical model, which simulates player movement to determine pitch control. This approach offers several advantages compared to the classical method. First, it provides a more accurate evaluation of pitch control for players and teams, allowing us to assess the significance of considering a uniform speed for all players. Second, it draws conclusions regarding how differences in player top speeds relate to their roles within teams, namely, goalkeepers, defenders, midfielders, and forwards. As we will see, the impact of player top speed is not uniform across these positions. Third, our model allows to simulate scenarios in which player top speeds are increased or decreased as a result of improvements or deteriorations in fitness. To this end, we introduce the *stamina factor*, a control parameter that adjusts player top speed and exhibits nonlinear behavior when correlated with pitch control. In the following sections, we define the numerical model and describe the datasets used to calculate pitch control. Next, we analyze the effects of top speed on players, teams, phases of the match (attack/defense), and halves of the match (first/second). Our analysis shows the importance of incorporating player top speed into pitch control models.

## Methodology Datasets

Pitch control models strongly depend on the data used for their calculation. In our case, data collection was performed using the Tracab Optical Tracking System, which employs a network of multi-camera units positioned around the stadium. This system captures the precise location of each player on the field at 25 frames per second,

with a spatial resolution of up to 10 cm<sup>25</sup>. Specifically, the system is based on a stereo multi-camera configuration comprising three units, each providing a resolution of 1920 × 1080 pixels. These units deliver a panoramic image that enables the triangulation of the location of the players and the ball. The accuracy of datasets generated by the Tracab Optical Tracking System has been validated through comparisons with GPS data in previous studies<sup>26,27</sup>. Together with the players' locations, the information about the team in possession at each frame is also included in the datasets, with a flag indicating if the ball is in-play. Concerning the players, their roles within teams are also considered and classified into four categories: goalkeepers, defenders, midfielders, and forwards. Substitutes have been excluded from the analysis, as their performance is constrained by shorter playing time. Under this framework, the tracking datasets containing the positional data of players were acquired from 100 matches of the first division of the Spanish National League (LaLiga) during the 2019/2020 season. Importantly, all matches analyzed were played prior to the outbreak of the COVID-19 pandemic, which led to a three-month suspension of the competition, affecting players' physical performance. Finally, the data used in this study has been previously anonymized and all methods were carried out in accordance with relevant guidelines and regulations. We obtained the approval from the Ethics Committee of the King Juan Carlos University (Madrid, Spain) to carry out this research (project PID2023-147827NB-I00), which has been carried out following all guidelines and regulations. The Spanish National League (LaLiga) approved to publish the results in an open access journal.

### The pitch control model

A pitch control model relies equally on the quality of the datasets and the modeling approach used in its design. The pitch control at a specific location is defined as the probability that a player or team will secure possession if the ball is sent directly there. Models of pitch control simulate both ball and player dynamics by considering their current positions and speeds. This means that effective ball control depends not only on proximity but also on the direction and velocity of the players; those moving swiftly toward the ball are likely to reach it more quickly than those approaching from a different direction who must first change their course.

To develop the pitch control model, a number of computations are required at each point on the field. These involve calculating the time needed for the ball to travel from its initial position to the target location, estimating how long it would take each player to get there, and then determining the overall probability that a team will gain control once both the ball and a player have arrived. Our approach is akin to the models proposed in<sup>23,28</sup>.

In our model, the ball is assumed to move at a constant speed of  $v_b = 15$  m/s. Consequently, the ball's arrival time at any location is given by

$$\tau_{\text{ball}} = \frac{\Delta x_b}{v_b},$$

where  $\Delta x_b$  is the Euclidean distance from the ball's starting position to the target location. Furthermore, as an initial assumption, all players are assigned a top speed of  $v_{\text{max},p} = 5$  m/s, which is considered a representative estimate of the top speed achievable when contesting for ball control. It should be noted that this uniform speed assumption is standard in classical pitch control models; however, the model will later be adapted to incorporate player-specific speed variations. To estimate a player's expected arrival time at a given location, denoted as  $\tau_{\text{player}}(\vec{r}, t_r)$  (with  $\vec{r}$  representing the displacement vector from the player's starting point to the destination and  $t_r$  the reaction time), we use a two-step approximation. First, a reaction time of  $t_r = 0.7$  seconds is assumed<sup>28</sup>; during this interval, the player is considered to continue along their current trajectory at the same speed, reaching a position  $\vec{r}_{\text{react}}$ . Once the reaction time elapses, the player is modeled as running directly toward the ball at their maximum speed  $v_{\text{max},p}$ . Finally, the player's total expected arrival time is computed by combining these two phases:

$$\tau_{\text{player}}(\vec{r}, t_r) = t_r + \frac{|\vec{r} - \vec{r}_{\text{react}}|}{v_{\text{max},p}}. \quad (1)$$

After determining the arrival times for both the ball and the players, the next step is to compute the likelihood that a player gains control of the ball at the intended target location. To achieve this, we adopt Spearman's assumption<sup>23</sup> that ball control is a stochastic process governed by an exponential distribution with a constant rate  $\lambda$ , which is the reciprocal of the mean time needed for a player to secure possession. Consequently, during any small time interval  $\Delta t$  in which a player is near the ball, the chance of controlling it is  $\lambda \cdot \Delta t$ . Consistent with<sup>23</sup>, we set  $\lambda = 4.3 \text{ s}^{-1}$  as our baseline control rate. Moreover, goalkeepers are assigned a higher control rate,  $\lambda_{\text{GK}} = 12.9 \text{ s}^{-1}$ , to reflect their enhanced ability to claim the ball, especially given their capacity to use their hands.

To further refine the model, an uncertainty parameter  $\sigma$  is introduced into the players' arrival times to account for factors not explicitly modeled. Combining these elements, the probability  $F_{\text{int},j}(\vec{r}, t, \sigma, t_r)$  that player  $j$  intercepts the ball at time  $t$  and at location  $\vec{r}$  is described by the cumulative distribution function of a sigmoid:

$$F_{\text{int},j}(\vec{r}, t, \sigma, t_r) = \frac{1}{1 + e^{-\frac{t - \tau_j(\vec{r}, t_r)}{\sqrt{3}\sigma/\pi}}}. \quad (2)$$

Note that this probability requires incorporating the actions or performance of other players on the field. Therefore, the overall probability that player  $j$  controls the ball when moving toward the location specified by the displacement vector  $\vec{r}$ , in the presence of other  $k$  players, is described by the following differential equation:

$$\frac{dPC_j}{dt}(t, \vec{r}, \sigma, \lambda_j, t_r) = \left(1 - \sum_k PC_k(t, \vec{r}, \sigma, \lambda_k)\right) F_{\text{int},j}(t, \vec{r}, \sigma, t_r) \lambda_j, \quad (3)$$

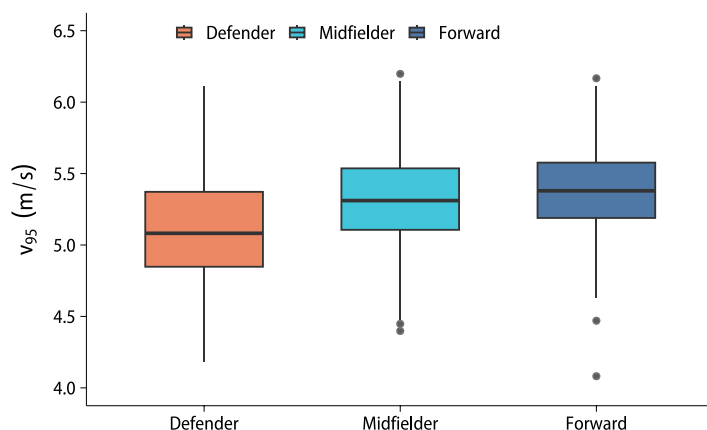
where  $PC_j$  denotes the potential Pitch Control field for player  $j$  and  $\lambda_j$  is the individual control rate. The term  $\sum_k PC_k(t, \vec{r}, \sigma, \lambda_k)$  represents the cumulative pitch control field of all other players  $k$  on the pitch at time  $t$ .

By integrating Eq. 3 over the interval  $t \in [t_{\text{ball}}, t_{\text{ball}} + 10]$  seconds and initializing  $PC_j(t, \vec{r}, \sigma, \lambda_j) = 0$  at the start, we obtain the control probability for each player at every point on the pitch. These values are then aggregated for all teammates to derive the overall pitch control for each team. Note that our model does account for differences between players due to their top speed (i.e., not all players are equal), in contrast to the model introduced in<sup>23</sup>, where differences between players depended only on the “initial conditions”, i.e., their position and current speed at the moment of the pitch control calculation.

### Speed dependence and the *stamina factor*

The next step is to consider the heterogeneity of the player speeds in the pitch control. Conventional pitch control models assign a uniform top speed  $v_{\text{max},p}$  to all players<sup>20,23</sup>, thereby simplifying computations at the expense of accuracy. To address this limitation, we first computed the probability distribution of each player  $i$ 's speed during every half of the match and extracted the 95th percentile threshold, denoted as  $v_{95}(i)$ . We then replaced the generic parameter  $v_{\text{max},p}$  with these player-specific values  $v_{95}(i)$ . This adjustment incorporates individual running capabilities into the model, allowing each player's pitch control to reflect their reported top speed. Note that, since pitch control is effectively a “zero-sum game”, an increase in one player's control necessarily results in a decrease in another's. We call the speed-dependent pitch control model as  $PC_v$ .

Moreover, the capacity to run faster does not imply that a player will always utilize their top speed. Depending on the match context or their physical condition, players may choose to conserve energy rather than run at full speed. Conversely, over-motivated players may increase their top speed during specific matches. This observation prompted us to explore the effects on pitch control when a player runs at a different pace than usual, either faster or slower. To capture this variability, we introduce a “stamina factor”  $\xi$  into the model, which multiplies the player's top speed (i.e.,  $v_{\text{max},p}(i) = \xi \cdot v_{95}(i)$ ), increasing it when  $\xi > 1$  and decreasing it when  $\xi < 1$ . We selected the 95<sup>th</sup> percentile as the reference for player top speed, which is commonly used in this type of analysis. Although other top speed values can also be considered (as long as the same reference is applied to all players), the 95<sup>th</sup> percentile velocity, also known as stride velocity 95th centile (SV95C), is a metric derived from wearable devices that represents the velocity at which a person walks or moves. It specifically identifies the speed below which 95% of a person's strides fall, effectively capturing the typical or near-maximal speed while excluding the fastest 5% of strides. Furthermore, this metric is often used in clinical trials<sup>29</sup>. Figure 1 shows the boxplots of  $v_{95}$ , in m/s, classified by playing position. As we can see, forwards exhibit the highest  $v_{95}$  values, followed by midfielders and then defenders, indicating positional differences in players' top running speeds. This visualization highlights the variability and distribution of the  $v_{95}$  speeds across playing positions and, therefore, suggests that considering the  $v_{95}$  may have consequences on the pitch control of players and teams.



**Fig. 1.** Boxplot of  $v_{95}$ , in m/s, by position: defenders (left), midfielders (center), and forwards (right). The  $v_{95}$  metric represents the 95<sup>th</sup> percentile of each player's running speed distribution. In the boxplot, the central line represents the median, the box boundaries correspond to the interquartile range (IQR), and the whiskers extend to the smallest and largest values within 1.5 times the IQR from the lower and upper quartiles, respectively. Individual points beyond the whiskers indicate potential outliers.

## Results

Once the methodology is defined, let us analyze how pitch control is related to the player top speed. Figure 2 shows an example of the classical pitch control calculated over a particular frame of a match. Colors (blue/red) indicate the team that would control the ball in case it reaches any region of the pitch and color intensity is proportional to the pitch control of each team. In the figure, we have included a vector indicating the direction of the speed of players, with a length proportional to the modulus. Players without an arrow are stationary. We can observe how the direction of the speed increases the probability of controlling the ball in the direction of the arrow. To elaborate this figure, the player speed is the one recorded at the specific frame of the figure, however, to simulate the ability of each player to move to other regions of the pitch to control the ball, all players are considered to run at the same speed, as it is assumed in classical pitch models. In the following sections, we will analyze the ability of players to move faster/slower than their counterparts.

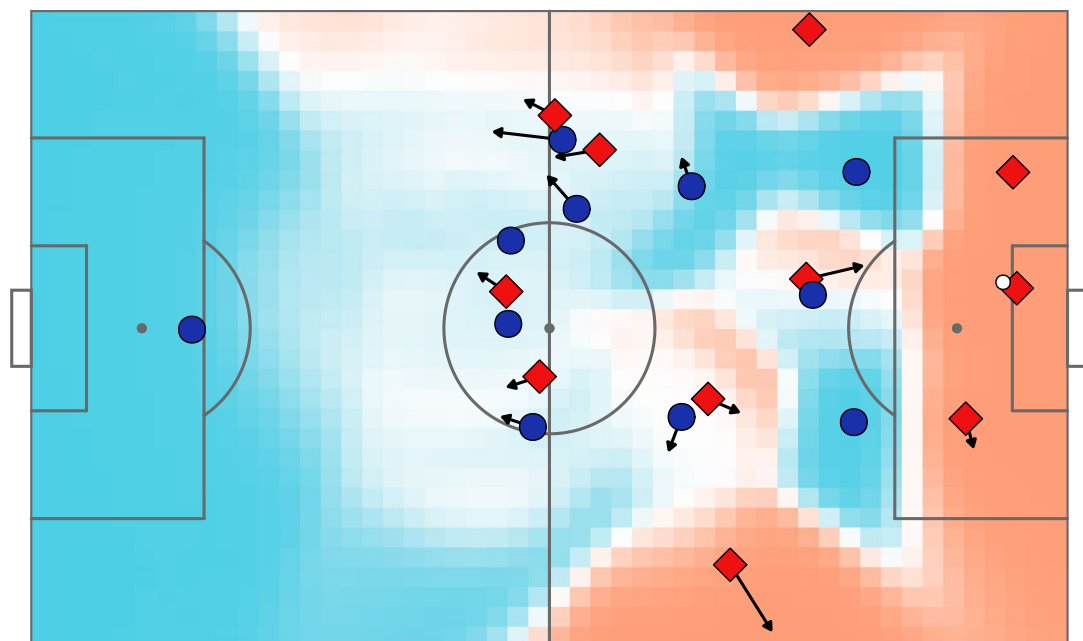
### Player pitch control vs player speed

As mentioned in previous sections, the fact that players' top speeds change from player to player has unavoidable consequences on their corresponding pitch control. Faster players are able to reach regions of the pitch than slower players when departing from similar distances. This fact is accounted for when considering the player's top speed  $v_{95}$  in a speed-dependent pitch control model  $PC_v$ . In Fig. 3 we show the percentage of pitch control accumulated by players in 90 effective minutes of play when their observed top speed is considered. In the figure, we have classified players into 3 different positions (defenders, midfielders, and forwards) and omitted the goalkeepers since their pitch control has different implications than the rest of the players. We can observe how the pitch control of players is positively correlated with their speed for all player positions (see correlations coefficients  $r$  at the figure). Interestingly, defenders, who are the players accumulating more pitch control, are the ones with the highest correlation coefficient. However, also note that they are not necessarily the players with the highest top speeds, indicating that not only the speed but also the role of the player, affects the values of the pitch control.

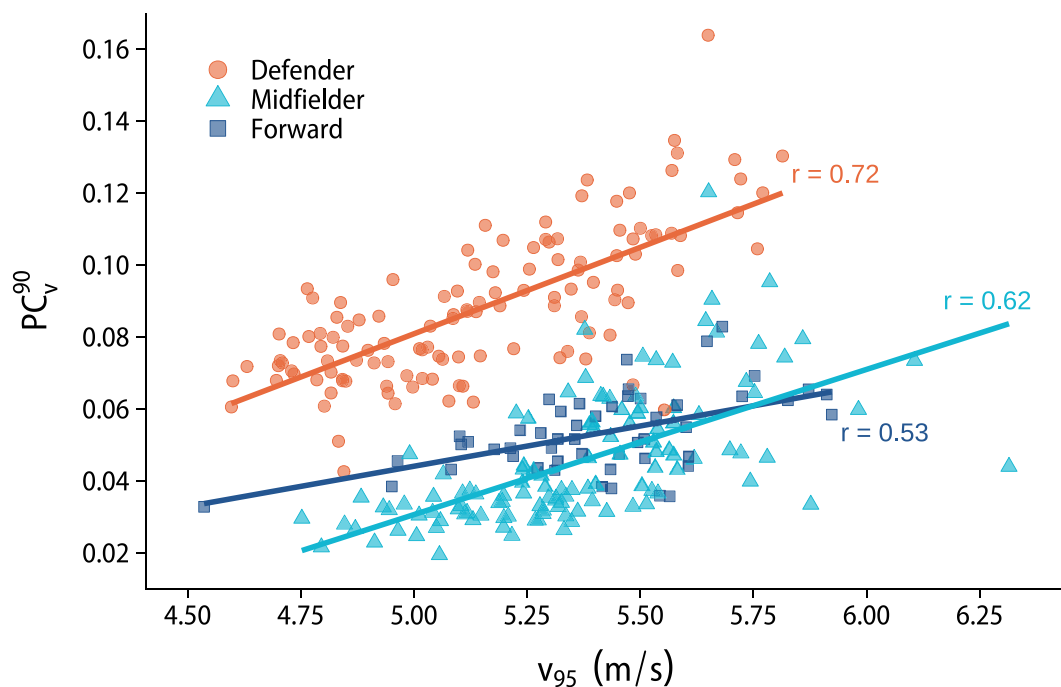
Figure 3 also highlights the fact that pitch control depends on more factors beyond the player top speed or the player position, as indicated by the low to moderate values of the correlation coefficients and the heterogeneity of the accumulated pitch control of players with similar top speeds.

### Team's pitch control

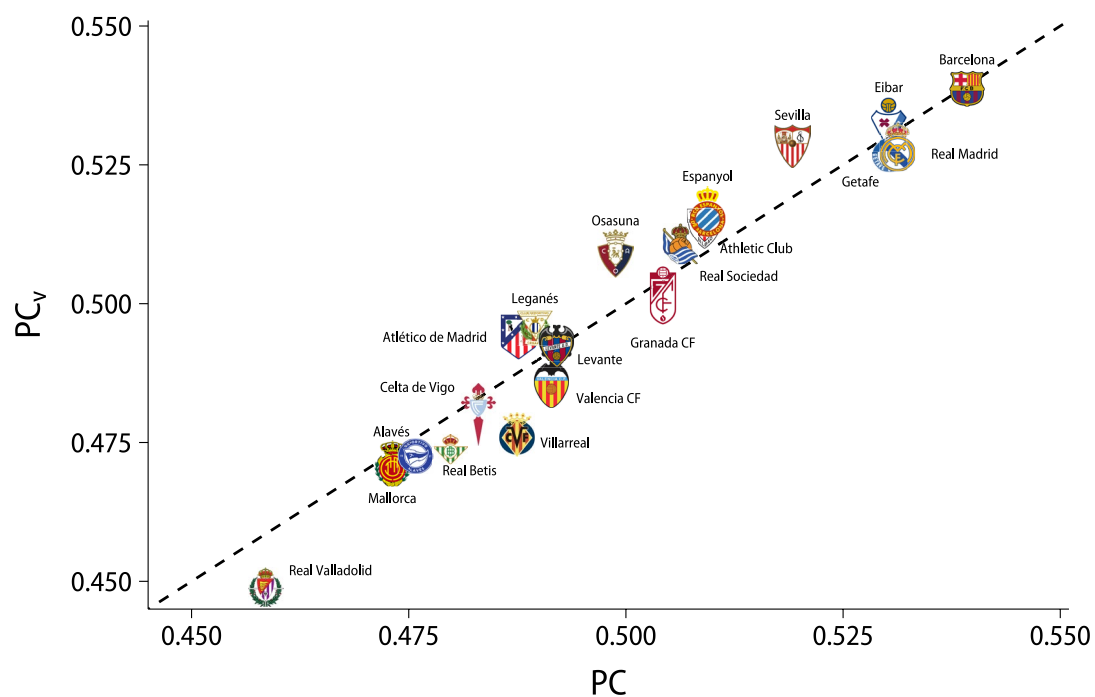
The consequences of including the player's top speed in the pitch control model can be projected into the team level. The fact that pitch control is a zero-sum game, where the increase of a given team results in the decrease of its rival, makes the analysis at the team level a useful tool to evaluate what teams are suffering the most from the physical performance of their players. With this aim, we analyzed how the pitch control of teams in the Spanish first division was affected by the inclusion of player top speed in the model. The pitch control of a team is defined as the sum of all team players' pitch control. Figure 4 shows a comparison of the classical pitch control ( $PC$ ) with



**Fig. 2.** Example of the pitch control (PC) calculated in a given frame of a match. Colors (blue/red) indicate the team with a higher probability of controlling the ball in case it moves to any region of the pitch. The length of the arrows assigned to each player is proportional to the speed at this specific frame. The player shape (circles or squares) indicates the team they belong to.



**Fig. 3.** Normalized speed-dependent pitch control  $PC_v^{90}$  as a function of the player top speed  $v_{95}$ . Each point corresponds to the pitch control accumulated by a player during 90 minutes. Three player positions are considered: defenders (circles), midfielders (triangles), and forwards (squares). Note the positive correlation  $r$  (Pearson correlation coefficient) in all positions.



**Fig. 4.** Classical pitch control ( $PC$ ) vs speed-dependent pitch control ( $PC_v$ ) for the Spanish teams at first division (season 2019/2020). Despite the positive correlation, there are deviations from  $PC(i) = PC_v(i)$ , which is indicated by the dashed line.



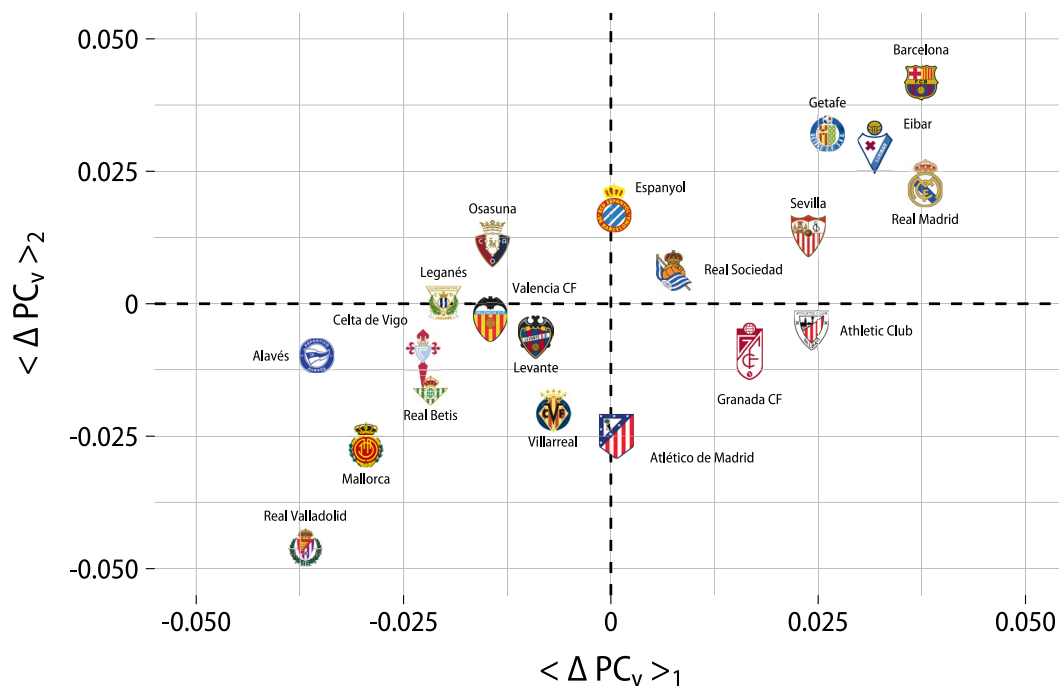
the speed-dependent pitch control ( $PC_v$ ) for each single team. Both variables are averaged over 10 matches. Note that teams with values over 0.5 accumulate more pitch control than their rivals.

Teams located above the identity line ( $PC_v(i) > PC(i)$ ) see their pitch control enhanced when player speeds are taken into account, while teams below the line experience a reduction in pitch control under the new model. The interpretation of the increase/decrease of the pitch control is not straightforward. It may reflect a playing style that emphasizes tactical organization or positional discipline over high-speed transitions, or simply differences in the physical profiles of the squad. It is important to note that pitch control is only one of many aspects that contribute to team performance on the field. It is worth noting how teams like Sevilla show both high pitch control and further improvements when speed is considered, suggesting a strong alignment between their tactical execution and physical capabilities. Notably, F.C. Barcelona ranks highest in both classical and speed-dependent pitch control, highlighting the consistency of their spatial dominance regardless of the modeling approach.

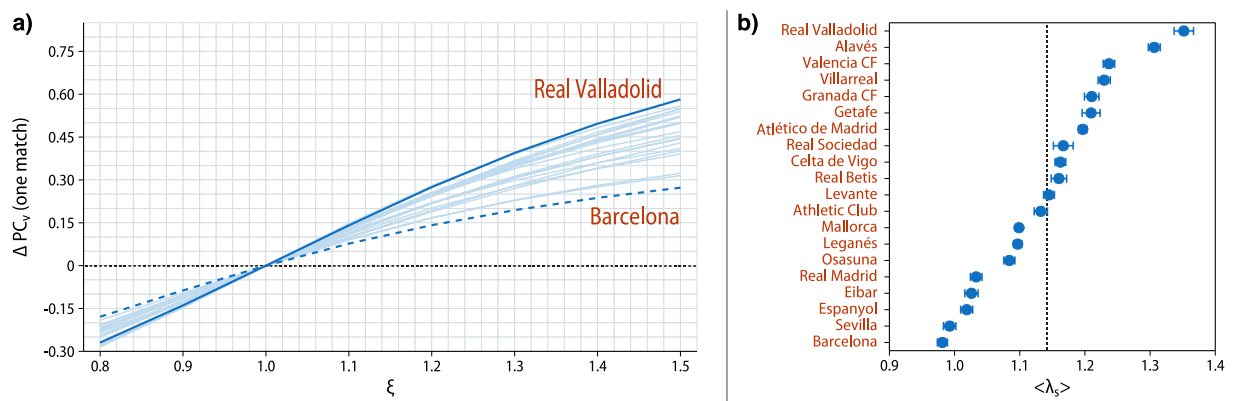
The fact that pitch control can be calculated at every frame allows to investigate its evolution. For example, in Fig. 5 we have split the analysis into the two halves of a match. For each team, we normalized the pitch control at each half of a match by subtracting the average of all teams  $\Delta PC_{vx} = PC_{vx} - \overline{PC}_{vx}$ , where  $x$  denotes the half. Then, we calculate the average over all matches of a team,  $\langle \Delta PC_v \rangle_{1,2}$ . This way, we are able to analyze what teams accumulated more pitch control than the others at each half of the matches. Figure 5 shows that, despite the positive correlation between the pitch control accumulated at both halves, not all teams behave in the same way. Comparing Fig. 4 and Fig. 5 we can observe how, for example, Getafe and Real Madrid have similar pitch controls; however, Getafe generates more pitch control in the second half than Real Madrid, whose pitch control is higher, on average, in the first half of a match. Another interesting case is the one of Atlético de Madrid, whose pitch control is average in the first half, but decays drastically in the second half.

### Stamina factor

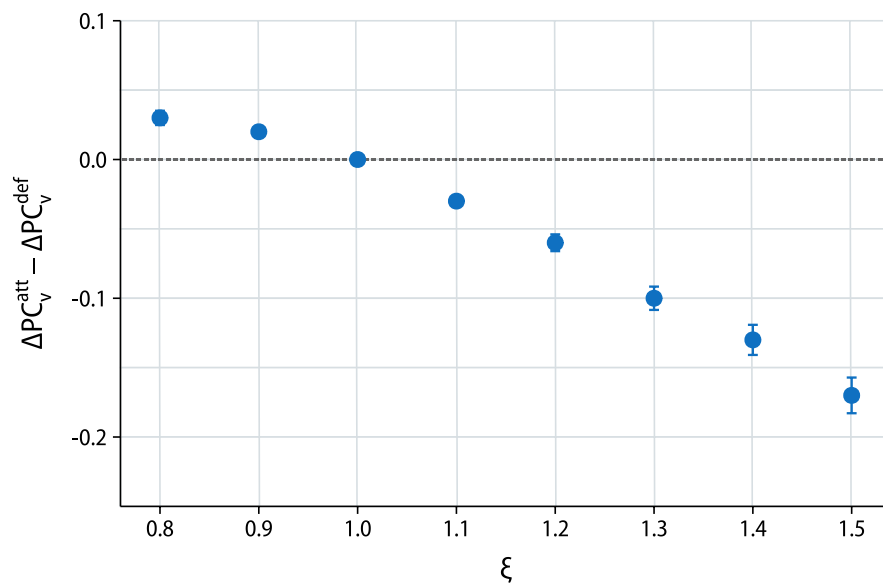
The next step is to evaluate what is the influence of increasing/decreasing player speed in the accumulated pitch control of teams. The stamina factor  $\xi$  modifies the top speed of each player and the pitch control can be recalculated with the modified speeds. This way, we can analyze how pitch control would be increased/decreased when a player can run faster/slower, thus evaluating the tentative impact of improving/worsening its fitness. Figure 6(a) shows the variation of pitch control  $\Delta PC_v$  for different values of the stamina factor for one match of each team. We can observe how all teams have similar behavior, which basically consists of a positive correlation that fits with a logarithmic function of the form  $f(x) = \lambda_s \log x + b$ . Then, we fit this function for the  $\Delta PC_v$  for each game and obtain the average scaling parameter  $\langle \lambda_s \rangle$  for each team. Note that the scaling parameter  $\lambda_s$  controls the dependence of the pitch control of a team on the stamina factor. Figure 6(b) shows the ranking of teams according to its corresponding  $\langle \lambda_s \rangle$ . As we will discuss later, teams with a lower  $\langle \lambda_s \rangle$  are the less affected by modifications of the player top speed, while teams with the highest values are those which would be benefited the most if the speed of its players was increased. However, they would also suffer more a decrease on top speed.



**Fig. 5.** Pitch control of all teams in the competition depending on the part of the match. The pitch control accumulated during the first and second halves,  $\langle \Delta PC_v \rangle_1$  and  $\langle \Delta PC_v \rangle_2$ , respectively, has been normalized by subtracting the average,  $\overline{PC}_{vx}$  of all teams at each half  $x$ . Dashed lines correspond to the averages of all teams at each half of the match.



**Fig. 6.** In (a), pitch control variation in a single match,  $\Delta PC_v$ , as a function of the stamina factor  $\xi$ , for each team. Each line corresponds to one of the twenty teams playing in the competition. We can observe how they follow a logarithmic distribution in all cases. In (b), we show the average scaling factor  $\langle \lambda_s \rangle$  of each team, ranked according to its value. The distribution indicates what teams would benefit the most from the stamina factor (i.e., Valencia CF, Alavés, and Real Valladolid). The vertical dashed line is the average of the scaling factor for the whole competition.

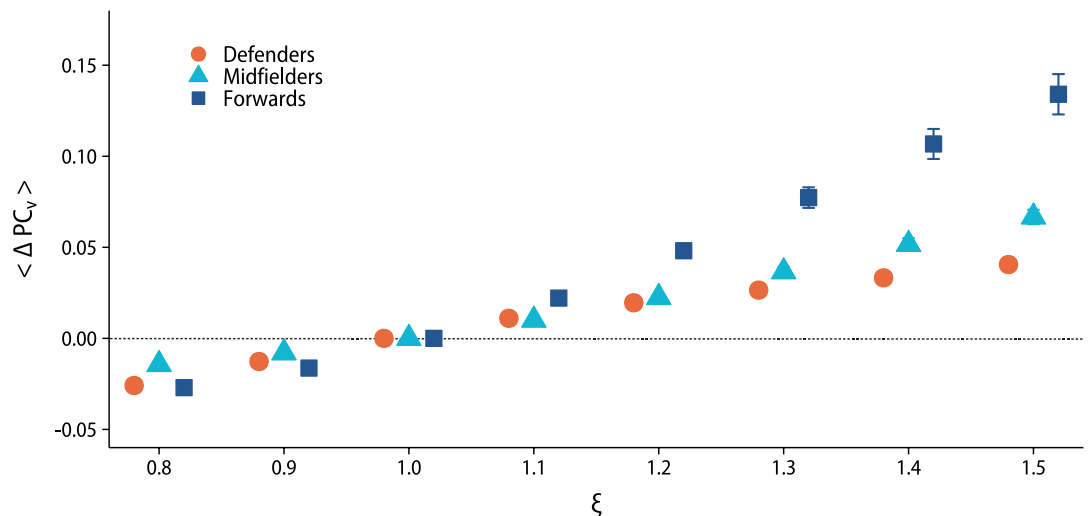


**Fig. 7.** Effect of the stamina factor during each match phase: attack vs. defense. Difference between the increment of pitch control when attacking ( $\Delta PC_v^{att}$ ) and defending ( $\Delta PC_v^{def}$ ), for multiple values of the stamina factor  $\xi$ . Points represent the average for each team in the competition, with error bars corresponding to one standard deviation.

Furthermore, different kinds of analysis can be performed by filtering the states of the match when the increase of pitch control is calculated. For example, in Fig. 7 we distinguish the impact of the stamina factor between the attacking state and the defending state of a team. Interestingly, while the impact of the stamina factor is positive in both cases, i.e., increases pitch control for  $\xi > 1$  and decreases it when  $\xi < 1$  (not shown), the amount of the impact is not the same. The subtraction  $\Delta PC_v^{att} - \Delta PC_v^{def}$  yields negative values for  $\xi > 1$  and positive values for  $\xi < 1$  indicating that  $\Delta PC_v^{def}$  is consistently larger in magnitude than that in  $\Delta PC_v^{att}$ . Note that for  $\xi < 1$  both are negative, but  $\Delta PC_v^{def}$  is more negative, indicating a major loss. Similarly, for  $\xi > 1$  both are positive, but  $\Delta PC_v^{def}$  is larger, indicating a larger benefit.

Finally, it is also possible to translate the analysis of the stamina factor to the level of players. In Fig. 8 we plot the increase of pitch control as a function of the stamina factor for the three main roles of the players: defenders, midfielders and forwards. We can observe how the forward players are the most affected by the positive increments of the stamina factor, followed by midfielders and defenders. However, there exists a certain degree of asymmetry in the positive/negative increments of the stamina factor. While the forward players are the ones that, on average, would benefit the most from an increase in their speed, a decrease produces much more





**Fig. 8.** Team average speed-dependent pitch control variation  $\langle \Delta PC_v \rangle$  vs the stamina factor  $\xi$ , for each role of a player. For each  $\xi$  and pair team/game, we only modify the top speed of the players with the given role. Then, we estimate the variation of the pitch control,  $\Delta PC_v$ , and normalize it by the number of players per modified position in that team and game. Lastly, we average the variations over all games played by a team for each role considered: defenders (circles), midfielders (triangles) and forwards (squares).

moderate differences. Two conclusions can be extracted from this fact. First, the effect of increasing/decreasing the top speed of players is not linear and, second, the role of the player in the team, probably constraining the location on the pitch, is strongly related with the effects of the stamina factor on the pitch control.

## Discussion

Integrating pitch control models into soccer performance analysis leads to a significant advancement in understanding both individual player dynamics and team strategies. These models offer a more detailed perspective on how players influence the available space on the field and shape offensive and defensive actions. However, player speed is a key determinant of a team's capacity to create or restrict space, thereby influencing strategic decisions and match outcomes. In that sense, we have shown that, by accounting for variations in player top speed, pitch control models yield more precise insights into a player's contribution to collective movement and spatial dominance. Incorporating each player's specific top speed reveals differences in individual performance with regard to the pitch control accumulated over a match. Interestingly, these differences are influenced by the player's role on the team; for example, defenders tend to accumulate higher pitch control (excluding goalkeepers), even though they are not necessarily the fastest players on the field. This observation underscores the complex interplay between individual top speed and pitch control. At the collective level, our analysis demonstrates that incorporating the actual top speed data recorded during matches affects the overall pitch control accumulated by soccer teams. While there is a positive correlation between the classical pitch control and the one that contains the player top speed, deviations from this trend emerge. For example, some teams, such as Sevilla F.C., Osasuna, and Atlético de Madrid, exhibit increased pitch control when player-specific top speeds are incorporated, whereas teams like Real Valladolid or Villarreal experience a decrease. This fact is not good or bad by itself, but indicates that not considering the player top speed leads to an inaccurate quantification of pitch-control that is specially significant in certain teams. Our results point to the importance of using actual top speed and suggest that more sophisticated models that include a “player avatar” (i.e., containing information of each player such as the top speed, acceleration or strength) could better characterize individual and team performance. It is also noteworthy that the impact of top speed on pitch control is not uniform throughout a match. This is reported when examining the effect of player top speed on pitch control separately for each half. Our analysis reveals that not all teams consistently increase or decrease their pitch control across both halves, indicating that this phenomenon is complex and phase-dependent. For instance, while FC Barcelona demonstrates above-average pitch control in both halves, Osasuna shows increased control in the second half. In the case of Granada, the highest pitch control is in the first half.

The observation that individual top speed influences pitch control opens the possibility of exploring how voluntary changes in top speed, whether due to motivational factors or improved physical conditioning, might affect pitch control. We observed that pitch control exhibits a logarithmic dependence on the stamina factor, a parameter that modulates player top speed. All teams increased (decreased) their accumulated pitch control for stamina factor values greater (less) than one. However, although the relationship is logarithmic for all teams, the scaling factor varies, indicating that some teams are more sensitive to changes in player top speed than others. Notably, teams such as Real Valladolid, Alavés, and Valencia CF would benefit the most from higher top speeds, whereas FC Barcelona, Sevilla, and Espanyol would experience smaller gains.

Another finding that highlights the complexity of the analysis is the different impact of the stamina factor during defensive versus offensive phases. Our results indicate that for positive values of  $\xi$ , the increase in pitch control is more pronounced during defensive phases than during attacking ones. This suggests that enhancing players top speed may be particularly beneficial for defensive efforts, a consideration that could be taken into account when programming the player physical loads during training. Beyond its methodological contribution, our approach of creating individualized player avatars, including parameters such as top speed, acceleration, and potential fatigue resistance, could have significant practical applications. For example, fitness monitoring data could dynamically update player avatars to reflect fatigue during matches, allowing coaches to make tactical adjustments based on anticipated declines in pitch control. Similarly, scouting and recruitment processes could use these avatars to identify players whose physical profiles best fit specific tactical systems, while performance analysts might apply pitch control scaling factors to assess how physical traits influence a team's spatial dominance across different phases of play. On that sense, future research should explore these applied uses to bridge the gap between theoretical models and practical decision-making. Moreover, the impact of the stamina factor appears to be closely tied to a player's role on the team. Forwards benefit the most from an increase in top speed, while defenders experience the smallest improvements in pitch control. This may be attributed to the distinct positional roles and movement patterns on the field: defenders typically maintain a steady pace, whereas forwards are more likely to engage in intermittent bursts of acceleration to outmaneuver opponents. Future research should focus on linking the outcomes of our methodology to the tactical organization of each team, as interpreting pitch control results requires detailed knowledge of team-specific playing styles and strategic decisions. Despite our study establishes a framework for assessing the influence of top speed on pitch control, we do not currently have access to data describing the tactical organization or specific tactical choices of the teams analyzed. Such information would be essential to explain, for instance, whether a team deliberately reduces pitch control in advanced areas to maintain defensive stability through a low-block formation. Therefore, applying our methodology in collaboration with coaching staff, who possess in-depth knowledge of their own tactical structures, would be crucial for achieving a more precise and meaningful tactical interpretation of the results. Another limitation of our study concerns the generalization of the results. We did not have access to competitions in different countries and, therefore, cannot claim that similar results would be obtained in competitions with different contexts.

In light of these findings, we believe that analyzing player speeds in relation to pitch control is an intriguing area with applications that extend beyond the scope of this paper. Future work could involve developing comprehensive player avatars that incorporate detailed information on physical and tactical performance, technical abilities, and, where feasible, psychological attributes. Such an approach would enable a better understanding of the complex behavior of soccer players and, more broadly, any team sport where spatiotemporal interactions between teammates and opponents are crucial. In addition to developing more comprehensive models for describing player performance, further analyses should focus on identifying the contextual variables that influence player behavior. On the one hand, the fact that a soccer match progresses through different phases suggests that a player's top speed may be related to these phases and affected by factors such as the match score (or score difference), match intensity, or specific tactical situations that increase or decrease player speed. On the other hand, contextual variables such as a team's ranking in the competition, or the importance of certain matches (due to rivalries or the necessity of winning) could also significantly influence a player's top speed and, consequently, a team's pitch control. As soccer continues to evolve toward data-driven analysis, pitch control models that consider the complexity of the problem and contain the player specific kinematics will be essential for advancing performance evaluation, tactical decision-making, and team coordination.

### Data availability

The data that support the findings of this study are available from LaLiga but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of LaLiga. You can contact Javier M. Buldu for any enquiry regarding the data.

### Code availability

The pitch control algorithm utilized to generate these results is accessible at: [https://github.com/Markfds01/Tracking\\_Laliga.git](https://github.com/Markfds01/Tracking_Laliga.git). Due to confidentiality agreements, the dataset used in this study cannot be publicly disclosed.

Received: 25 April 2025; Accepted: 30 September 2025

Published online: 05 November 2025

### References

1. Okubo, A. Dynamical aspects of animal grouping: swarms, schools, flocks, and herds. *Adv. Biophys.* **22**, 1–94 (1986).
2. Sumpter, D. J. The principles of collective animal behaviour. *Philos. Trans. R. Soc. B: Biol. Sci.* **361**(1465), 5–22 (2006).
3. Vicsek, T. & Zafeiris, A. Collective motion. *Phys. Rep.* **517**(3–4), 71–140 (2010).
4. Hamilton, W. D. Geometry for the selfish herd. *J. Theor. Biol.* **31**(2), 295–311 (1971).
5. Vicsek, T., Czirok, A., Ben-Jacob, E., Cohen, I. & Shochet, O. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.* **75**(6), 1226 (1995).
6. Gudmundsson, J. & Horton, M. Spatio-temporal analysis of team sports. *ACM Comput. Surv.* **50**, 22 (2017).
7. Balague, N., Torrents, C., Hristovski, R., Davids, K. & Araújo, D. Overview of complex systems in sport. *J. Syst. Sci. Complex.* **26**, 4–13 (2013).
8. Buldú, J. M., Gómez, M. A., Herrera-Diestra, J. L. & Martínez, J. H. Nonlinear dynamics and networks in sports. *Chaos Solit. Fractals* **142**, 110518 (2021).

9. Memmert, D., Lemmink, K. A. & Sampaio, J. Current approaches to tactical performance analyses in soccer using position data. *Sports Med.* **47**(1), 1–10 (2017).
10. Cervone, D., D'Amour, A., Bornn, L. & Goldsberry, K. A multiresolution stochastic process model for predicting basketball possession outcomes. *J. Am. Stat. Assoc.* **111**(514), 585–599. <https://doi.org/10.1080/01621459.2016.1141685> (2016).
11. Seidl, T., Cherukumudi, A., Hartnett, A., Carr, P. & Lucey, P. Bhostgusters: Realtime interactive play sketching with synthesized NBA defenses. In **Proceeding of the 12th MIT Sloan Sports Analytics Conference**, Boston, MA. Boston: MIT (2018).
12. Bridgeman, L. A. & Gill, N. D. The use of global positioning and accelerometer systems in age-grade and senior rugby union: a systematic review. *Sports Med. Open* **7**, 1–34 (2021).
13. Rahimian, P. & Toka, L. Optical tracking in team sports: A survey on player and ball tracking methods in soccer and other team sports. *J. Quant. Anal. Sports* **18**(1), 35–57. <https://doi.org/10.1515/jqas-2020-0088> (2022).
14. Torres-Ronda, L., Beanland, E., Whitehead, S., Sweeting, A. & Clubb, J. Tracking systems in team sports: a narrative review of applications of the data and sport specific analysis. *Sports Med. Open* **8**(1), 15 (2022).
15. Kovalchik, S. A. Player tracking data in sports. *Annu. Rev. Stat. Appl.* **10**(1), 677–697 (2023).
16. Novillo, Á. et al. Analysis of player speed and angle toward the ball in soccer. *Sci. Rep.* **14**(1), 11780 (2024).
17. Marcelino, R. et al. Collective movement analysis reveals coordination tactics of team players in soccer matches. *Chaos Solit. Fractals* **138**, 109831 (2020).
18. Buldú, J. M. et al. Soccer tracking networks: Beyond event-based connectivity. In *Conference on Analytics in Sports Tomorrow 2020*, ed F.C. Barcelona., 1–13 (2020).
19. Chacoma, A., Billoni, O. V. & Kuperman, M. N. Complexity emerges in measures of the marking dynamics in soccer games. *Phys. Rev. E* **106**(4), 044308 (2022).
20. Taki, T. & Hasegawa, J. I. Dominant region: a basic feature for group motion analysis and its application to teamwork evaluation in soccer games. In *Videometrics VI*, 3641,48–57, S.P.I.E. (1998).
21. Voronoi, G. M. Nouvelles applications des parametres continus a la theorie des formes quadratiques. *J. Reine Angew. Math.* **134**, 198 (1908).
22. Kim, S. Voronoi analysis of a soccer game. *Nonlinear Anal.: Model. Control* **9**(3), 233–240 (2004).
23. Spearman, W., Basye, A., Dick, G., Hotovy, R. & Pop, P. Physics-based modeling of pass probabilities in soccer. In *Proceeding of the 11th MIT Sloan Sports Analytics Conference* (2017).
24. Tuo, Q., Wang, L., Huang, G., Zhang, H. & Liu, H. Running performance of soccer players during matches in the 2018 FIFA World Cup: Differences among confederations. *Front. Psychol.* **10**, 1044 (2019).
25. <https://tracab.com/products/tracab-technologies/tracab-optical/> [Accessed on 10th May 2024].
26. Pons, E. et al. A comparison of a GPS device and a multi-camera video technology during official soccer matches: Agreement between systems. *PLoS One* **14**(8), e0220729 (2019).
27. Felipe, J. L., Garcia-Unanue, J., Viejo-Romero, D., Navandar, A. & Sánchez-Sánchez, J. Validation of a video-based performance analysis system (Mediacoach\*) to analyze the physical demands during matches in LaLiga. *Sensors* **19**(19), 4113 (2019).
28. <https://github.com/Friends-of-Tracking-Data-FoTD/LaurieOnTracking> [Accessed on 10th January 2024].
29. Servais, L. et al. Stride velocity 95th centile: insights into gaining regulatory qualification of the first wearable-derived digital endpoint for use in Duchenne muscular dystrophy trials. *J. Neuromuscul. Dis.* **9**(2), 335–346 (2022).

## Acknowledgements

J.M.B. acknowledges the support by Ministerio de Ciencia e Innovación (Spain) under grants PID2020-113737GB-I00 and PID2023-147827NB-I00. A.A. acknowledges support from the Grant No. RYC2021-033226-I funded by MICIU/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR. Y.M. was partially supported by the Government of Aragón, Spain, and “ERDF A way of making Europe” through Grant No. E36-23R (FENOL). A.A. and Y.M. acknowledge the support by Ministerio de Ciencia Innovación y Universidades, Agencia Española de Investigación (MICIU/AEI/10.13039/501100011033) Grant No. PID2023-149409NB-I00. M.F. acknowledges the support by the Government of Aragón through a Government of Aragón Ph.D. contract order ECU/592/2024.

## Author contributions

JMB, AA, MFdS, YM and AN: study conception and design. JMB, AA and MFdS: experimental design. RLdC and RR: data collection. All authors: analysis and interpretation of results. JMB, AA and MFdS: draft manuscript preparation. All authors reviewed the results and approved the final version of the manuscript.

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to J.M.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025