

Quantifying Off-Ball Defensive Impact through Cover Shadows

Lorenzo Cascioli^{1,*}, Allen Wang^{1,*}, Luca Stradiotti¹, Maaïke Van Roy¹, Pieter Robberechts¹, Maxim Wouters², Arne Jaspers^{1,2} and Jesse Davis¹

¹ *KU Leuven, Dept. of Computer Science; Leuven.AI, B-3000 Leuven, Belgium*

² *Royal Belgian Football Association, RBFA Knowledge Centre, B-1480 Tubize, Belgium*

** These authors contributed equally.*

Introduction

Defensive contributions in soccer often occur without the ball and are poorly captured by traditional metrics. In fact, Wu and Swartz [1] observe that players typically have possession for less than three minutes per match. Yet, almost all of the widely reported metrics are based on actions involving the ball. In practice, a defender might merely block an opponent's passing lane: such a move might prevent a valuable pass that is never realized and hence, never recorded. These off-ball actions, often referred to as "cover shadows", are crucial to a team's defense but have received little analytical attention.

Existing defensive metrics fail to capture this positional aspect of defending. Conventional stats (tackles, interceptions, clearances) only record events where the ball is won or play is interrupted, ignoring instances where defenders merely steer attackers into less dangerous options. More sophisticated frameworks (i.e., expected disruption models) [2], [3], aim to value off-ball defensive actions by attributing changes in a team's typical playing style or offensive efficiency to the defensive behavior of opponents. However, these approaches mainly consider man-marking actions and primarily attribute their impact at the team level. In short, no existing approach explicitly quantifies the impact of a defender obstructing an opponent's passing lane.

Recent data advances make new analyses possible. StatsBomb 360 data (launched in 2021) provides the locations of all players surrounding each event, effectively combining traditional event logs with tracking-like spatial context. This "contextual" event dataset provides the kind of context needed to evaluate defensive positioning. By leveraging these positional snapshots, one can identify when a defender's placement cuts off a passing lane and model how much the opponent's attack is thereby disrupted.

In this paper, we propose a novel data-driven approach to quantify the effect of blocking passing lanes in soccer. First, we determine whether a player blocks a passing lane. Naively, this means checking if the player is on an imaginary line between the ball and a potential receiver. In practice, however, the ability to intercept also depends on factors beyond positioning such as reaction time, ball speed, and trajectory. To capture this, we propose a physics-based model that estimates the

probability of a player blocking a pass at any location by incorporating time to intercept (cf. pitch control) and an approximation of the ball's travel time given the target location of a potential pass.

Second, we measure how effectively a defending team is blocking passing lanes. We start by quantifying the threat of the attacking team in a game snapshot. To do so, we sum the threat of all potential passes, weighting each pass by its likelihood of completion. We then create a counterfactual scenario by removing (one or more) defenders that are casting a cover shadow, and re-take the threat of this alternative game state. The difference in threat serves as a measure of cover shadows effectiveness. We show how this metric can be used in practice to analyze defensive contributions at both the player and team levels, thereby identifying those who excel or underperform in the use of cover shadows.

Finally, we assess if and how a team can improve their ability to block passing lanes. To do so, we cast defensive positioning as an optimization problem. The objective is to identify the optimal allocation of defenders to passing lanes so that the most dangerous attackers are blocked.

In short, this work makes three key contributions:

- **RQ1:** We provide a more accurate analytical method for identifying when defenders obstruct a passing lane.
- **RQ2:** We provide a novel metric for measuring the cover-shadow effectiveness of players and teams. Our metric rates how defenders' positioning influences the likelihood of opponents playing valuable passes.
- **RQ3:** We introduce a novel approach for optimizing defensive positioning that enables quantifying how small positional adjustments in a team's collective shape can reduce the attacking threat by blocking access to progressive passing lanes.

Data and Definitions

Dataset Overview

Our work builds on StatsBomb 360 event stream data, which augments the conventional human-annotated event logs with spatiotemporal snapshots extracted from broadcast video. Each event record includes the standard event metadata (i.e., event type, timestamp, player, team, bodypart, outcome, and pitch coordinates) along with a "freeze frame" capturing the positions of the players (where visible in the broadcast video) at the moment of the event, along with their relationship to the ball carrier (i.e., teammate or opponent). Because freeze frames occur at the moment of the ball-related action, they do not represent continuous player movement.

Our main dataset comprises 1,752 matches from the Big 5 European leagues during the 2024/25 season. We use the data from the Bundesliga, Premier League, LaLiga and Ligue 1 to train models and select hyperparameters. Serie A data is reserved to evaluate the models and develop the use cases. Since the StatsBomb 360 freeze frame data does not provide the identity of players, we complement this evaluation dataset with one match of the 2020 European Championship (Belgium

vs Portugal). For this match, we enriched the 360 snapshots with player IDs gathered from tracking data provided by the RBFA. The tracking data and StatsBomb 360 event stream data were first synchronized using ETSY [4]. Afterward, the player identities in each freeze frame were determined using a shortest-distance approach. This combined dataset will allow us to study defensive behavior at both the player and team level.

To standardize and facilitate analysis, we employ the SPADL¹ representation of the events and transform the player coordinates in the freeze frames into the 105 x 68 meters standardized pitch coordinate system. We exclude passes where the origin or the destination falls outside of the visible area, as we might be missing essential context.

Defining a (Blocked) Passing Lane

To analyze pass blocking, we select all snapshots associated with a pass in the attacking half where the ball carrier had at least one viable progressive passing option. We exclude set pieces, high or aerial passes, and passes made with body parts other than the foot. Passes that occur inside the penalty area are also removed. This process results in 67,426 snapshots from the Serie A season and 224 snapshots from the Belgium–Portugal match. Each snapshot provides the positions of players visible in the frame (not all 22 players) including:

- The ball carrier p_a
- All visible potential receivers $R = \{r_1, r_2, \dots, r_n\}$, and
- All visible defenders $D = \{d_1, d_2, \dots, d_n\}$

For each potential pass $\{p_a, r_i\}$ in the frame, we compute whether the passing lane is open or blocked given the positions of D . Therefore, we define a passing lane as the direct, straight-line corridor between the ball carrier and a potential receiver. Formally, the lane $L(p_a, r_i)$ is represented by the segment connecting their positions. A lane is considered blocked if any defender $d_j \in D$ is positioned such that they could realistically intercept the ball if the pass was attempted. We operationalize this definition in RQ1.

Identifying Lane-Blocking Players

Not all defenders contribute equally to blocking passing lanes. We distinguish between two primary defensive roles: man-marking defenders and lane-blocking defenders.

- **Man-marking defenders** closely follow an attacker and are not really concerned with blocking passing lanes.

¹ SPADL unifies existing event stream formats into a common language. It provides a tabular representation in which each on-the-ball action is described by the same set of twelve attributes. For more info, see <https://github.com/ML-KULeuven/socceraction>.

- **Lane-blocking defenders**, in contrast, occupy the empty space between the ball carrier and one or more potential receivers to discourage dangerous passes.

Operationally, we classify as man-markers all defenders that are positioned within a 3m radius centered 1m behind an attacking player, to account for the typical marking position where the defender stays close to the attacker between the latter and the goal. Defenders outside this marking zone are treated as potential lane-blocking defenders. In this work, we focus specifically on the contribution of these lane-blocking defenders.

RQ1: Is a player blocking a passing lane?

In previous work, quantifying whether a defender blocks a pass has been approached with several geometric approaches [5], [6], [7]. For example, the most common approach involves drawing a fictional cone extending from the passer to the receiver. A pass is then considered to be blocked if a defender is positioned within the cone. These geometric approaches provide a useful starting point for assessing blocked passes, but they overlook that players are already in motion before the pass is played and that several uncertain or unobserved factors (such as reaction time) render passing-lane obstruction a probabilistic event. Therefore, we introduce *Lane Control*, a novel physics-based pass block model.

Baseline Pass Block Models

We first define three geometric baseline models for evaluating if a pass is blocked. Figure 1 provides a graphical illustration of each approach.

First, the **Line Corridor** baseline assumes the passing lane to be a rectangle $R_w(p_a, r_i)$ of fixed width w , centered on the segment connecting the ball carrier p_a and receiver r_i :

$$R_w(p_a, r_i) = \{x \in \mathbb{R}^2 \mid \text{perp_dist}(x, L(p_a, r_i)) < w/2\},$$

where $\text{perp_dist}(x, L(p_a, r_i))$ represent the perpendicular distance from a point x to the line segment $L(p_a, r_i)$. A pass is considered blocked if any defender's position falls within this corridor:

$$\exists d_j \in D : d_j \in R_w(p_a, r_i).$$

We empirically obtained the best results for $w = 2$ meters.

Second, the **Cone Corridor** baseline expands the passing corridor to a cone with vertex at the ball carrier and base at the receiver. The cone width at the receiver scales proportionally with the pass length: $w = k * \|r_i - p_a\|$ where k is a scaling constant and $\|r_i - p_a\|$ is the Euclidean distance between r_i and p_a . Let $T_k(p_a, r_i)$ represent the cone triangle with vertex at p_a , base at r_i , and width scaling constant k . A pass is blocked if any defender lies within the cone:

$$\exists d_j \in D : d_j \in T_k(p_a, r_i).$$

In our experiments, we set $k = 0.2$ so the cone width is 2 meters for a 10 meter pass.

Finally, the **Gaussian Cone** baseline incorporates positional uncertainty about the defenders. Each defender is represented as an isotropic 2D Gaussian [8] $N(\mu_j, \Sigma_j)$ centered on their recorded position $\mu_j = (x_j, y_j)$ with covariance $\Sigma_j = \sigma^2 I$. For a pass from p_a to r_i , each defender is projected onto the pass center line; at that location, we take the cone's perpendicular cross section and compute the normalized Gaussian mass of the defender within the cross-section (the blocking fraction F_σ). A pass is considered blocked if any defender's blocking fraction F_σ exceeds a threshold τ :

$$\exists d_j \in D : F_\sigma(p_a, r_i, d_j) \geq \tau$$

All mathematical details and derivations of F_σ are provided in Appendix A. We use $\sigma = 1$ for 1m² variance, $\tau = 0.5$ for the threshold, and the same $k = 0.2$ so the cone width is 2 meters for a 10 meter pass.

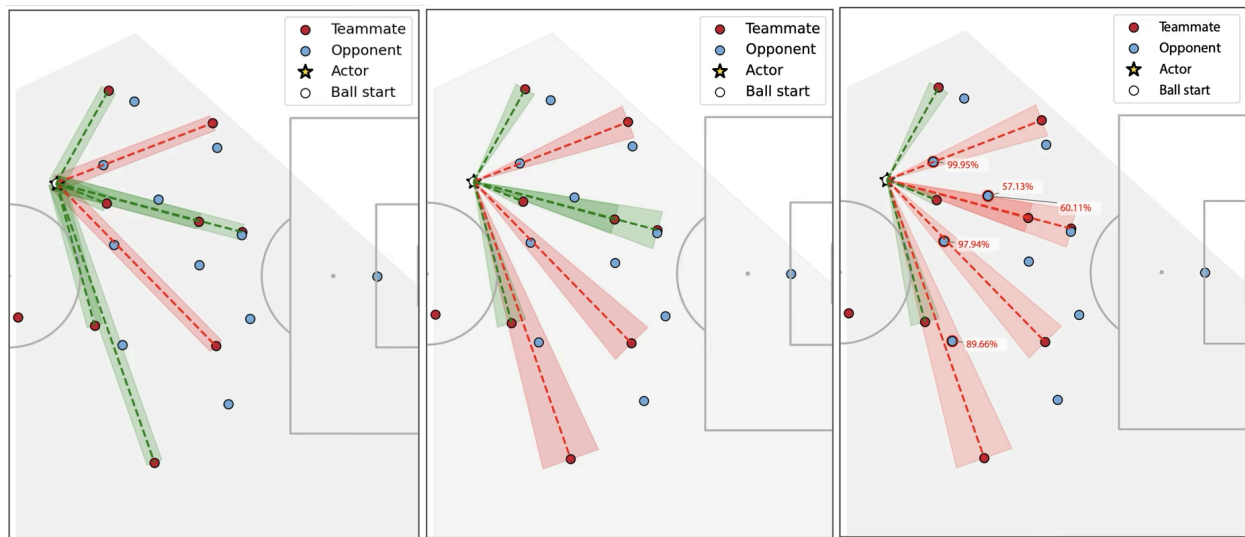


Figure 1: Sample illustration of blocked/open pass options in a game state with Line Corridor, Cone Corridor, and Gaussian Cone baselines. Green passes are marked as open, red passes are marked as blocked. Note that when considering the attacker closest to the attacking penalty box, the two closest defenders are both slightly outside of the line and the cone, as shown for the first two baselines. In this case, the Gaussian Cone still marks the passing lane as blocked due to the influence of the defender positioned earlier on the passing lane. The second defender is still ignored, as he is positioned right below the attacker, therefore he does not generate any valid cross section.

Because of their static assumptions, these baselines ignore how the movement of players and the ball shapes passing outcomes. They can identify whether a defender is geometrically positioned to block a pass, but not how players' initial speeds, their timing, or the ball's velocity influence the likelihood of interception.

Our solution: Lane Control

To address the limitations of the geometric baseline approaches, we introduce Lane Control, a physics-based model inspired by pitch control [9], [10], and pressing intensity [11]. In short, Lane Control evaluates the probability that a pass will be intercepted by computing the time-to-intercept (TTI) for the ball, attacking players, and defenders at a set of discrete locations describing the passing lane. Specifically, it consists of three subsequent steps: corridor parameterization, TTI estimation, and probability conversion.

Corridor Parameterization

First, Lane Control parameterizes the passing corridor into three lines, analogous to the previously defined passing cones: a center line corresponding to the shortest straight-line distance from the passer to the receiver; and a left and right line along the edges of the passing cone, ending at the two sides of the receiver. Concretely, each line is sampled at 30 target points from the passer to the receiver.

TTI Estimation

Having computed a discrete set of locations along the passing lane, we compute the TTI for the ball, all defenders D , and all attackers R to each of these locations. This is simply the estimated time that it takes the ball/player to get to that point, given the current velocity and constants describing the motion.

- Ball Time:** To model the ball's time, we follow Spearman's approach [9] for ball motion under quadratic drag, ignoring gravity and Magnus force while assuming a constant drag coefficient and a straight-line path. We use the same constants as the original model: air density ($\sigma = 1.22 \text{ kg/m}^3$), drag coefficient ($C_D = 0.25$), cross sectional area of the ball ($A = 0.038 \text{ m}^2$), and mass of the ball ($m = 0.42 \text{ kg}$). We combine these into a single drag parameter $k = (\sigma C_D A) / (2m)$. Then, the time for the ball to reach a target point at distance d given initial speed v_0 is defined as:

$$T_{ball} = \frac{e^{kd} - 1}{k v_0}.$$

In practice, the initial speed v_0 is often unknown or difficult to estimate. To address this, we set v_0 to an average ball speed of 12.0 m/s as a proxy. For numerical stability when kd is small, we use the numpy function `exp1` to compute $e^{kd} - 1$.

- Players TTI:** For each sampled target point q_k on any of the three lines and for each player $p \in D \cup R$, we model the TTI with a capped accelerate-then-cruise model with reaction time. Let r_p and v_p be the initial position and velocity of the player. After a reaction delay of

$t_{react} = 0.7$ s, the player is at $r_{react} = r_p + v_p t_{react}$. Let $d = \|q_k - r_{react}\|$ be the distance from the player to the target after the reaction and $\hat{e} = (q_k - r_{react}) / \|q_k - r_{react}\|$ be the unit vector from the player to the target. We only consider the initial velocity component towards the target: $v_0 = \max(0, v_p \cdot \hat{e})$. The maximum acceleration ($a = 7.0$ m/s²) and the maximum speed ($v_{max} = 12.0$ m/s) are fixed for all players. Defenders only need to get within a block radius ($r_{block} = 0.7$ m) of the target to block/deflect the pass, while receivers must reach the point exactly to properly control the ball. The effective distance is thus $d_{eff} = \max(d - r_{block}, 0)$ for defenders and $d_{eff} = d$ for attackers. Each player's time to reach maximum speed during the acceleration phase is $t_{accel} = (v_{max} - v_0) / a$ while the distance covered during acceleration is $d_{accel} = v_0 t_{accel} + \frac{1}{2} a t_{accel}^2$. Adding the reaction delay, the TTI for a player is then given by the piecewise formula below:

$$TTI_p(d_{eff}) = \begin{cases} t_{react} + \frac{d_{eff}}{v_{max}}, & \text{if } v_0 \geq v_{max}, \\ t_{react} + \frac{-v_0 + \sqrt{v_0^2 + 2a d_{eff}}}{a}, & \text{if } d_{eff} \leq d_{accel}, \\ t_{react} + t_{accel} + \frac{d_{eff} - d_{accel}}{v_{max}}, & \text{otherwise.} \end{cases}$$

In other words, a player can reach the target point in one of three cases:

- Cruising case: the player is already moving at maximum speed and simply has to react before covering the effective distance.
- Acceleration case: the player is not already running at max speed, hence he starts accelerating and reaches the target before hitting max speed.
- Accelerate-then-cruise case: the player accelerates, reaches max speed, and continues running at max speed before arriving at the target.

If a player is already running at max speed, then the TTI does not consider any acceleration. If instead the player gets to the ball before reaching the max speed, then max speed cruising is not considered. Some popular physics-based implementations do not consider the acceleration phase and simply assume that players run at max speed [12]. We found this to clearly worsen the method performance, hence we chose to explicitly model acceleration.

The presented formulation assumes we know the speed of each player in the snapshot. However, StatsBomb 360 data do not provide any velocity information. We therefore infer velocities at the time of the snapshot [13]. For each event, we look at the next event and calculate the positional displacement for all players in the event freeze frame and divide it by the difference in event timestamps to estimate velocities for each snapshot.

Probability Conversion

In the previous steps, we parametrized the passing line with a finite set of points, and calculated the TTI of all players and the ball to each of these points. As the final step, we now compute the probability that each of the three passing lines (right, center, left) is blocked by aggregating TTIs.

Similarly to [9], we express the probability that each player gains possession of the ball at a particular target point (k) as the combination of two components. First, a player needs to be in a position to intercept the ball at that location. We call the time it takes the ball to get there T_k . Then, we model the **interception probability** as the CDF of a sigmoid centered on $\Delta t = T_k - TTI_j$: $P_j^{int} = 1 / (1 + \exp(-\Delta t/s))$, with $s = \sqrt{3} \sigma / \pi$. This accounts for variability that is not explicitly modeled in TTI calculation such as differing speeds, reaction time, or fatigue. So, for example, if the player would arrive way before the ball, $\Delta t \ll 0$ hence $P_j^{int} \approx 1$. We found $\sigma = 0.20$ to give us the best results. Second, even if the player is ready, controlling the ball takes time. This **control probability** is modeled as an exponential process $P_j^{ctrl} = 1 - \exp(-\lambda \Delta T) \approx \lambda \Delta T_k$ with rate $\lambda \approx 4.3 \text{ s}^{-1}$ and $\Delta T_k = T_k - T_{k-1}$ (with T_0 the start) being the time interval between consecutive target points.

Combined, this gives the probability that player j gains possession in the interval $[T_k, T_k + \Delta T_k]$:

$$P_j(T_k, T_k + \Delta T_k) = P_j^{int}(T_k) \cdot \lambda \Delta T_k$$

This holds if only one player exists. In practise, we need to subtract the probability that any player has already intercepted the ball by time T . So the formulation becomes:

$$P_j(T_k, T_k + \Delta T_k) = (1 - \sum_l P_l(T_k)) \cdot P_j^{int}(T_k) \cdot \lambda \cdot \Delta T_k$$

For each passing lane, we compute this discretized numerical integration simultaneously for each player, across each $[T_k, T_k + \Delta T_k]$ interval along the parameterized line. This gives us the total probability P_j that each player intercepts/receives the ball before the end of the cone:

$$P_j^{total} = \sum_{k=1}^K \left(1 - \sum_l P_l(T_k) \right) \cdot P_j^{int}(T_k) \cdot \lambda \cdot \Delta T_k$$

We aggregate the probabilities of players on the defending team into a block probability $P_{blocked}$ and those of the attacking team into a receive probability $P_{received}$. Note that as we don't run the

numerical integration to $t \rightarrow \infty$, we are left with a non-zero probability that no one controls the ball before the end of the cone.

We define any one of the three lines as blocked if $P_{blocked} > P_{received}$. Finally, the passing lane can be defined as blocked if **Any** (one), **Majority** (two), or **All** (three) of these lines are blocked. This provides more granularity and flexibility to our analysis. We give a visual example of the discretized passing lanes, an intuition about TTIs and final blocking/receiving probabilities in Figure 2.

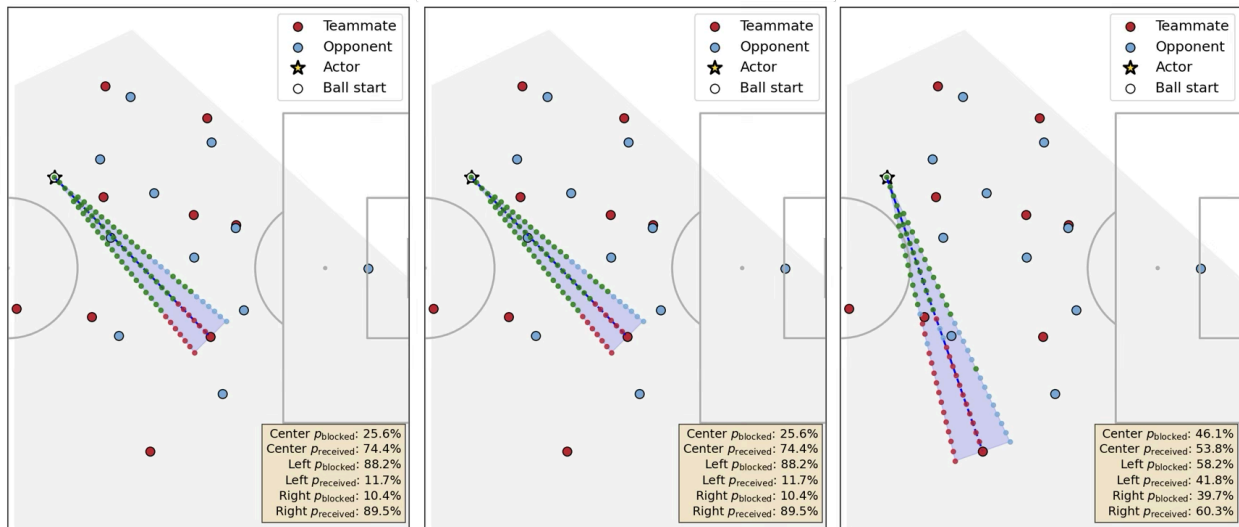


Figure 2: Evaluating block probability for potential passes using Lane Control. The dots along the passing cone represent the target points, and are colored according to the object that has the shortest time to arrive: the ball (green), an attacker (red), and a defender (blue). For each of the center, left, and right lines of the passing corridor we show the probability that a pass along that line is blocked $P_{blocked}$ and that it is successfully received $P_{received}$.

Evaluating Pass Block Models

Evaluating the accuracy of our Lane Control approach presents a fundamental challenge: by definition, a blocked pass is an unobserved event. When a defender successfully obstructs a passing lane, the corresponding pass is typically never attempted, meaning there is no explicit ground truth record against which to validate the prediction of $P_{blocked}$ and $P_{received}$. Consequently, assessing the quality of a pass lane-blocking model requires an indirect approach.

To overcome this issue, we evaluate our pass block models on observed passes. The key intuition behind our approach is that if a passing lane is genuinely blocked, it should not be possible for a successful pass to be played through it. Therefore, we assess the accuracy of our proposed Lane Control approach and the three geometric baselines by testing whether their predictions of blocked or open lanes are consistent with the results of passes that were actually executed during matches.

Formally, for each completed pass (p_a, r_i) , we compute whether the lane $L(p_a, r_i)$ was predicted to be blocked or open by each model. If a model predicts the lane as blocked but the pass was nevertheless successfully completed, this constitutes a false positive, indicating an overestimation of defensive influence. Conversely, if a model predicts the lane as open but the pass was unsuccessful or intercepted, this is treated as a false negative, representing an underestimation of defensive influence.

We restrict our evaluation to progressive passes (excluding high passes) which occurred between the halfway line and the attacking penalty box, as our original intuition was to evaluate the impact of defensive midfielders screening the backline. Although this evaluation could be extended to the entire pitch and consider the defensive impact of attackers pressing, we found expanding the area of interest made the already imbalanced dataset even more skewed. For progressive passes between the halfway line and attacking penalty box, the pass completion ratio was 98% across the Big 5 European leagues (Appendix B). This is mostly because we exclude high passes, which bypass the effect of blocking defenders while also accounting for a large percentage of incomplete passes. For example, ~73% (141/193) of incomplete passes in the match between Hellas Verona and Napoli on August 18th, 2024 were high passes.

Table 1 reports standard classification metrics across the Big 5 European leagues. Both Lane Control Majority and All outperform the three baselines in terms of balanced accuracy (+ 11-15%) and macro F1 score (+ 0.17-0.20). While all of the models had similarly low rates of false negatives (underestimating defenders influence on a pass that was actually unsuccessful), the baselines had significantly higher rates of false positives (overestimating defenders influence on a pass which was actually completed). The raw accuracies are high due to the class imbalance. Lane Control Majority has slightly better recall (hence the higher balanced accuracy), while Lane Control All has lower overestimation (FP) and higher precision (hence the best F1/macro F1 score).

Table 1: Performance of baseline pass block models and Lane Control on progressive passes between midfield and attacking penalty box for Big 5 European leagues from 2024/25 season (Ground truth: 69,500 passes analyzed, 98.74% successful passes, 1.17% unsuccessful passes)

	Metric					
	FP (overestimation)	FN (underestimation)	Balanced Accuracy	Accuracy	Macro F1 Score	F1 Score
Line Corridor	10.89% (7,572)	0.90% (628)	55.82%	88.20%	0.3006	0.0430
Cone Corridor	2.17% (1,508)	1.07% (745)	53.03%	96.76%	0.2932	0.0561
Gaussian Cone	14.09% (9,791)	0.84% (586)	56.79%	85.07%	0.3048	0.0417

Lane Control (Majority)	1.52% (1,053)	0.73% (507)	68.01%	97.76%	0.4806	0.2811
Lane Control (All)	0.98% (680)	0.76% (529)	66.93%	98.26%	0.4941	0.3189

RQ2. How effective are the cover shadows?

Having established how to detect when a defender blocks a potential pass, we now extend our focus to the collective level. Our aim is to quantify how effectively the defending team’s positioning constrains the opponent’s most dangerous, chance-creating passes. This entails adding an additional layer to the analysis: the threat brought by each potential pass.

Ideally, a set of well-executed cover shadows reduces the threat that the ball carrier can generate by passing to teammates. Therefore, to quantify how effective the defending team’s cover shadows are, we:

1. **Quantify the current offensive threat of the attacking team** as the sum of each available passing option’s threat value, weighted by the probability that the ball gets to that player (i.e., the ball carrier chooses that option for the next pass, and the pass succeeds).
2. **Quantify the threat of a counterfactual unblocked scenario** by modeling how pass selection and success probabilities would change if fewer or no cover shadows were applied by the defending team.
3. **Define the blocking score** as the difference between the two threat values, where a positive value indicates that the applied cover shadows are actively reducing the opponents’ potential to create goal-scoring chances.

Quantifying the Current Threat

The threat posed by the attacking team in the current snapshot S is composed of the individual threat of each of its players:

$$threat_S = \sum_{r \in R} threat(r | S)$$

Intuitively the threat posed by each of the attacking team’s players depends on two factors: how dangerous each attacking player’s position is, and how likely that player is to successfully receive the ball next. An attacker may occupy a dangerous area (e.g., inside the box) but is unlikely to receive the ball if surrounded by several defenders. In this case, the ball carrier might choose a safer pass to an open teammate, or the attempted pass may fail as it is missed or intercepted by a defender. Consequently, the most dangerous attackers are those who combine a high positional threat (i.e., proximity to goal) with a high probability of receiving the ball (i.e., being relatively open).

It follows that computing the attacking team's threat requires estimating (1) the positional threat of each potential receiver, and (2) the probability that the given receiver will successfully receive the next pass.

Passes often target the open space near a player rather than their exact location. Therefore, both aspects should not only consider the receiver's exact location, but also his surroundings. To do so, we discretize the pitch into a 104x68 cell grid, where each cell is described by its (x, y) coordinates, and we consider that the ball carrier can attempt a pass towards each of these cells. We then partition the space computing Voronoi regions so that each grid cell is associated with the closest attacker. We express the attacker controlling position (x, y) as:²

$$r^*(x, y) = \arg \min_r \text{dist}((x_r, y_r), (x, y)).$$

Defining the set of cells assigned to receiver r as N_r , the threat associated with receiver r in snapshot S can be computed by taking the weighted sum over all $(x, y) \in N_r$:

$$\text{threat}(r | S) = \sum_{(x,y): r^*(x,y)=r} \text{threat}(x, y) \cdot p_r(x, y)$$

where $\text{threat}(x, y)$ represents the positional threat at location (x, y) , and $p_r(x, y)$ represents the probability that the given receiver r successfully receives the ball at location (x, y) .

We estimate the probability that receiver r successfully receives the ball at location (x, y) as

$$p_r(x, y) = P(D = (x, y) | S) \cdot P(o = \text{success} | D = (x, y), S),$$

where:

- $P(D = (x, y) | S)$ is the probability that the next pass targets cell (x, y)
- $P(o = \text{success} | D = (x, y), S)$ is the probability that if attempted, a pass to cell (x, y) succeeds.

Therefore the formula for $p_r(x, y)$ quantifies the probability that the given receiver will successfully receive the next pass by encompassing (1) pass selection probability, i.e., how likely it is that the receiver is chosen for the next pass and (2) pass success probability, i.e., how likely it is that a pass to that receiver succeeds (and hence is not missed or intercepted by a defender). We now explain how we estimate the positional threat, the pass selection likelihood and the pass success likelihood.

Positional Threat Model

² This simplification assigns each cell to one, and only one, attacker. However, it would also be possible to apply a probabilistic assignment.

To estimate the positional threat for every cell $(x, y) \in N_r$, we employ Expected Threat (xT) [14].³ This approach overlays a grid on the field (Figure 3) and assigns a value to each grid cell based on how often a possession from that area leads to a goal.

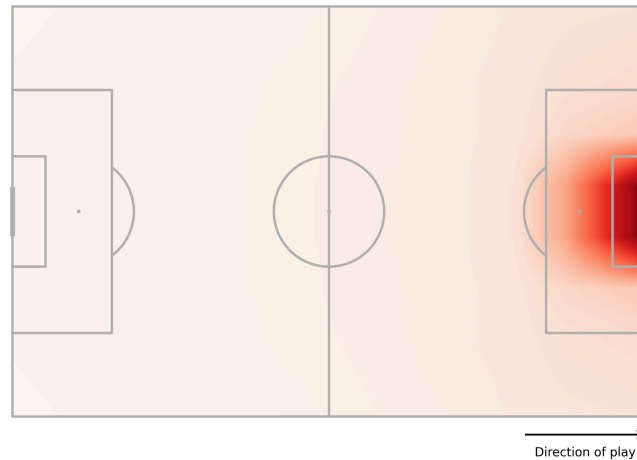


Figure 3: The positional threat as a result of upscaling the original Expected Threat (xT) grid. Darker red means a higher positional threat (i.e., probability of scoring a goal later on in the possession).

Pass Selection Model

To estimate the pass selection probabilities, we take inspiration from previous work [6] and implement a pass selection model using the SoccerMap [15] architecture. This popular CNN architecture has been used (among other things) to generate a probability surface over the full pitch, estimating the probability that the ball carrier will play the ball at each location. To do so, it takes into account the positional information of all players on the pitch, as these will all influence the probability a pass is played to a particular location.

We train our pass selection model by gathering all passes from the 2024/25 season of the Big Five European leagues, filtering as described in "Data and Definitions". We use four leagues for training, extracting a total of 1,116,738 passes. A random sample of 20% of these passes is used as a validation set for model selection, and the rest is used as the training set. We keep aside the 317,061 passes from Serie A, as we use them for evaluating the model. Figure 4 (left) shows the predicted pass selection surface for one example pass.

Pass Success Model

³ We use Karun Singh's league-wide xT values from the 2017-18 Premier League season (available from https://karun.in/blog/data/open_xt_12x8_v1.json), upscaled to a 104x68 grid with bilinear interpolation.

Similarly, we use SoccerMap to generate a probability surface over the full pitch, estimating the probability that the pass will be successfully played to each location. We again reproduce the model from [6], and use the same set of passes as for the pass selection model. The difference is in the input (we only use the first 7 channels for pass success) and in the label, which reflects whether an executed pass succeeded or not.

The pass success surface for the example snapshot is shown in Figure 4 (right). We precisely evaluate our pass selection and pass success models and discuss their performances in Appendix C.

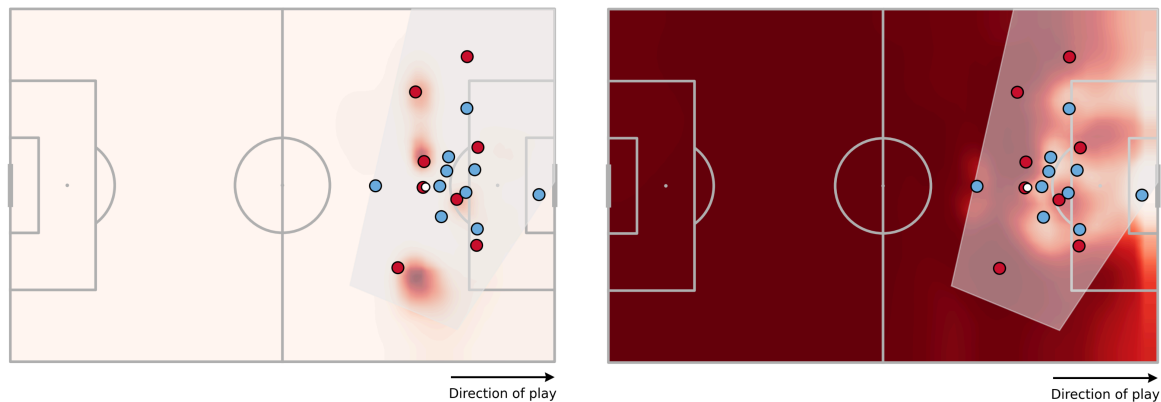


Figure 4: Pass selection surface (left) and pass success surface (right) for an example snapshot. The pass selection surface shows how likely it is that the ball carrier plays a pass to each location. Red shading represents more likely pass destinations. The pass success surface shows the probability that a pass to any location successfully reaches the receiver. Red shading represents easier passes.

Ultimately, the product of the xT surface (Figure 3), the pass selection surface (Figure 4, left) and the pass success surface (Figure 4, right) yields the weighted threat surface shown in Figure 5. Each cell in the resulting surface amounts to $threat(x, y) \cdot p_r(x, y)$. We can finally aggregate threat values per each potential receiver r by taking the sum over all (x, y) : $r^*(x, y) = r$. Aggregated values are also reported in Figure 5.

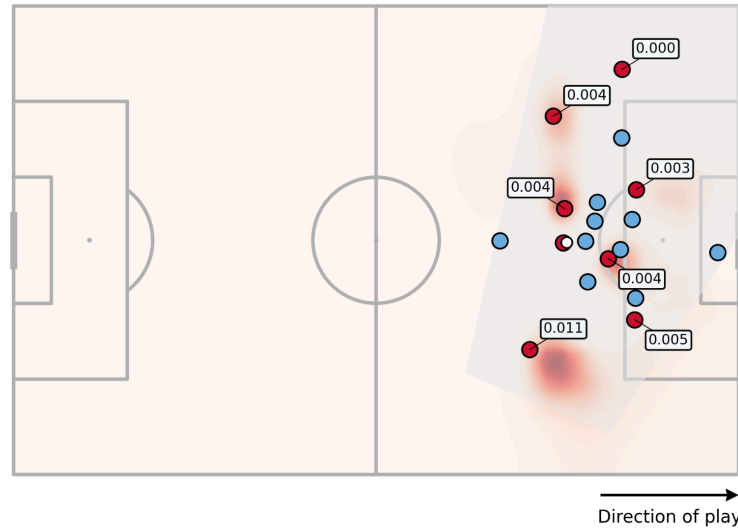


Figure 5: Weighted threat surface for the example snapshot, obtained as the product of xT surface (Figure 3), pass selection surface (Figure 4, left) and pass success surface (Figure 4, right). Red shading represents more threatening pass options. The threat of each player is the sum of the values of all cells in the Voronoi region of that player. The total offensive threat is the sum of the threat of each potential receiver, and is equal to 0.031.

Quantifying the Unblocked Threat

The task of lane-blocking defenders is to occupy dangerous passing lanes and thereby discourage passes to attacking players who are positioned in a threatening area. While the positional threat of a potential receiver is exclusively related to his location, the choice between the receivers and the difficulty of each potential pass are influenced by the defenders' positioning. Hence, the threat of the attacking team in a counterfactual *unblocked* scenario (S') can be estimated by modifying the defenders' positioning.

Specifically, we alter the input to the SoccerMap models by removing (a subset of) the defenders that are not man-marking. Depending on the use case, one can focus on removing one defender (e.g., to value individual covering abilities), or a coherent subset / all lane-blocking defenders (e.g., to quantify the covering abilities on a group level). Afterward, we estimate new pass selection and pass probability surfaces and recompute the total threat with updated weights $p_r(x, y)$ for every location. This ultimately yields a different threat value for the *unblocked* counterfactual scenario. We show the comparison between threat surfaces of the original and counterfactual scenario in Figure 6, together with aggregated threat per player.

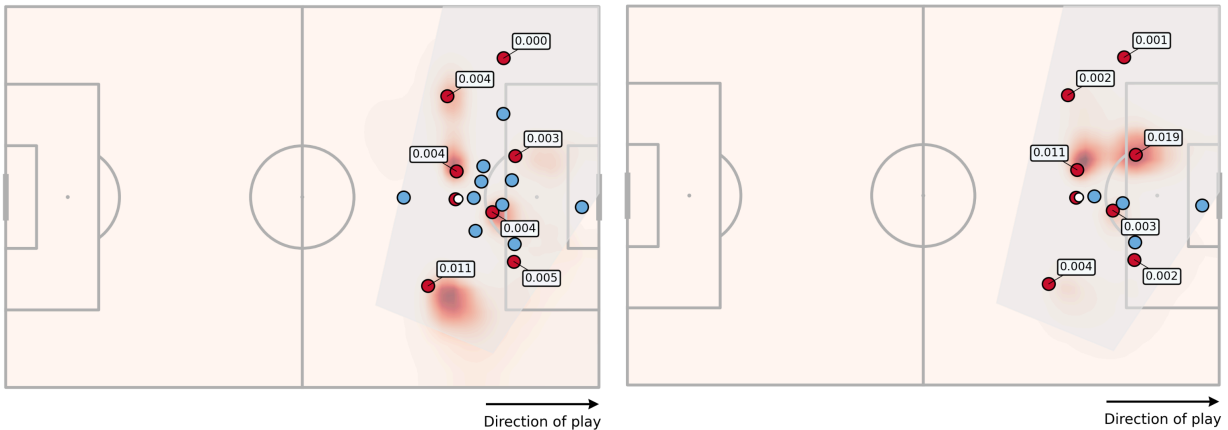


Figure 6: Weighted threat surface for our example snapshot in the original (left) and counterfactual scenario where cover shadows are removed (right). Only defenders that are man-marking are kept in the counterfactual. Red shading represents more threatening pass options. The total threat of the original scenario is 0.031, and increases to 0.041 in the counterfactual, yielding a blocking score of 0.01. A positive blocking score indicates that cover shadows are preventing some threat.

As a second example, we show that we can also study the effect of a single cover shadow by only removing one player in the counterfactual. Figure 7 focuses on the midfielder of the defending team (in purple), who is shielding a dangerous pass. If we ignore him in the figure on the right, the attacking team's striker is by far the most threatening receiver, as he is well-positioned and has some free space to occupy in proximity of the box. On the left, we see that the cover shadow applied by the defender is a well-executed block forcing the ball carrier to play a safer pass (e.g., on the left wing).

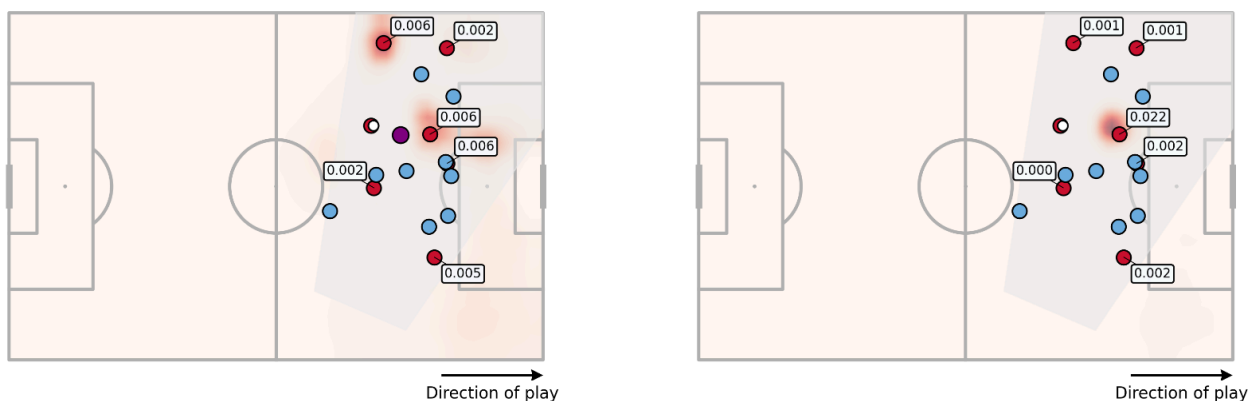


Figure 7: Threat surfaces for a game snapshot where we study the effect of a single cover shadow. We focus on the purple defender. On the right, we see that removing him clearly increases the threat coming from the attacker he is blocking.

Defining the Blocking Score

To quantify the overall offensive threat the defending team's cover shadows are preventing, we define the blocking score as:

$$\text{blocking_score} = \text{threat}_{s'} - \text{threat}_s$$

In our running example from Figures 6, the defensive team's cover shadows clearly impact the aggregated threat values of the two alternatives. Indeed, $\text{threat}_{s'} = 0.041$ and $\text{threat}_s = 0.031$, hence the blocking score is equal to 0.01. The positive value indicates that the well-executed cover shadows have reduced the offensive team's danger by 25%. In the second scenario (Figure 7) the players are further away from the goal, hence the offensive team is less threatening, with $\text{threat}_{s'} = 0.027$ and $\text{threat}_s = 0.029$. The blocking score is still positive, albeit much lower (0.002).

RQ3: Can the cover shadows be improved?

Finally, we study if and how the effectiveness of the employed cover shadows can be improved. To this end, we compute the optimal defensive positioning of the defending team with respect to the most dangerous attackers. We formulate this task as a linear optimization problem.

Intuitively, we attempt to find the minimal distance each defender has to travel such that the passing lanes to the most dangerous attackers are blocked. This requires defining (1) which attackers are considered the most dangerous, and (2) in which areas of pitch the defenders should be positioned in order to guarantee optimal blockage of the aforementioned attackers. Then, it becomes possible to optimize the defenders' positions by assigning each individual defender to a location within the considered areas minimizing displacement from their current positions. Searching for the minimal distance to travel allows us to find *realistic* adjustments to the defenders' positioning in order to optimize the effectiveness of their cover shadows.

Defining the set of dangerous attackers

We define the set of dangerous attackers to consider in the optimization problem, $R' \subset R$, as those attackers that are positioned between the ball carrier and the goal. We quantify how dangerous each attacker r_i is by exclusively considering the xT value at his location. We ignore the likelihood term from RQ2 to make the problem formulation more tractable when repositioning defenders.

Defining the set of areas that defenders should occupy

Ideally, defenders should block all passing lanes to dangerous attackers. We define an attacker to be blocked if at least one defender is positioned within the passing cone $T(p_a, r_j)$ going from the

ball carrier p_a to the attacker r_j (i.e., we use the cone baseline from RQ1 to define positional adjustments using straightforward geometric regions).

In a naive scenario where each defender can block at most one attacker, the set of areas to occupy simply corresponds to the cones related to the n_{def} attackers with the highest xT , where n_{def} equals the number of defenders in the snapshot. However, when passing cones overlap, a single defender could block multiple passing lanes at once when positioned at the intersection of these cones (see Figure 8). Therefore we consider additional areas of interest for defending positioning beyond the set of cones. For any combination of dangerous attackers $U \subseteq R'$, we define the intersection of their cone triangles as $\cap \{r_j \in U\} T(p_a, r_j)$. If this intersection is non-empty, we consider it as a potential area for a defender to occupy. To encompass all passing cone triangles and non-empty intersections that defenders could occupy, we refer to each of the considered areas as a polygon. The set of all considered polygons is formally defined as:

$$P = \{ \cap \{r_j \in U\} T(p_a, r_j) : U \subseteq R', \text{area}(\cap \{r_j \in U\} T(p_a, r_j)) > 0 \}$$

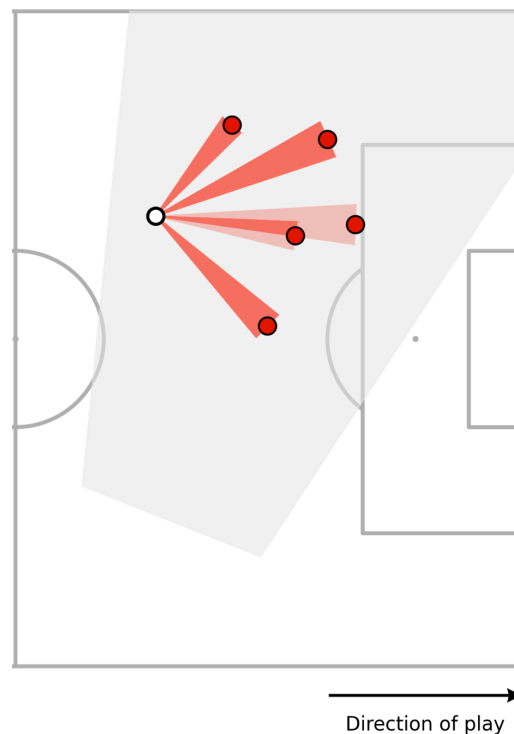


Figure 8: A snapshot with 5 dangerous attackers and their passing cones. Suppose $n_{def} = 4$. Whenever two cones intersect, it is enough for a defender to be positioned in the intersection of the two cones to simultaneously block two attackers. The dark red polygons show the areas in which 4 defenders could be positioned to block all 5 dangerous attackers. If we did not consider intersections of cones, we would have 5 cones and 4 defenders, hence the attacker with the lowest xT would remain free even under optimal defensive positioning.

We then need to identify the optimal subset of polygons, i.e., the smallest set of polygons that, when all occupied by defenders, guarantees the best possible coverage of dangerous attackers. To do so, we enumerate all possible combinations of polygons. For each subset of polygons Q , we check which attackers are covered, and we call this set $covered(Q)$. We compute the total xT blocked by Q as $\sum_{r_j \in covered(Q)} xT_j$. Having enumerated all possible subsets, we formally define the minimal-size set of polygons maximizing threat coverage as:

$$G = \operatorname{argmax}\{Q \subseteq P, |Q| \leq n_{def}\} \sum_{r_j \in covered(Q)} xT_j$$

Therefore, the best-case scenario for the defenders is to occupy each polygon in G . Figure 8 illustrates this in an example with 5 attackers for which the optimal set of polygons has size 4.

Defining the optimization problem

Once we know where defenders should be placed to optimally prevent passes to dangerous attackers, we can frame an optimization problem to suggest repositioning. Formally, the optimization problem then becomes:

$$\operatorname{minimize} \sum_{d \in D} \sum_{g \in G} c_{d,g} \cdot x_{d,g}$$

where D is the set of defenders and G the optimal set of polygons to occupy. Here, $c_{d,g} \in C$ (i.e., the cost matrix) represents the shortest distance that defender d needs to travel with respect to his current position in order to occupy polygon g (i.e., be inside polygon g). $x_{d,g}$ is a binary variable set to 1 if defender d is assigned to polygon g . We solve this linear optimization problem using the Hungarian algorithm [16], [17]. The solution to this problem yields updated coordinates for defenders that were not optimally positioned. Figure 9 shows an example situation in which the approach suggests slightly repositioning three defenders to better shield the attackers.

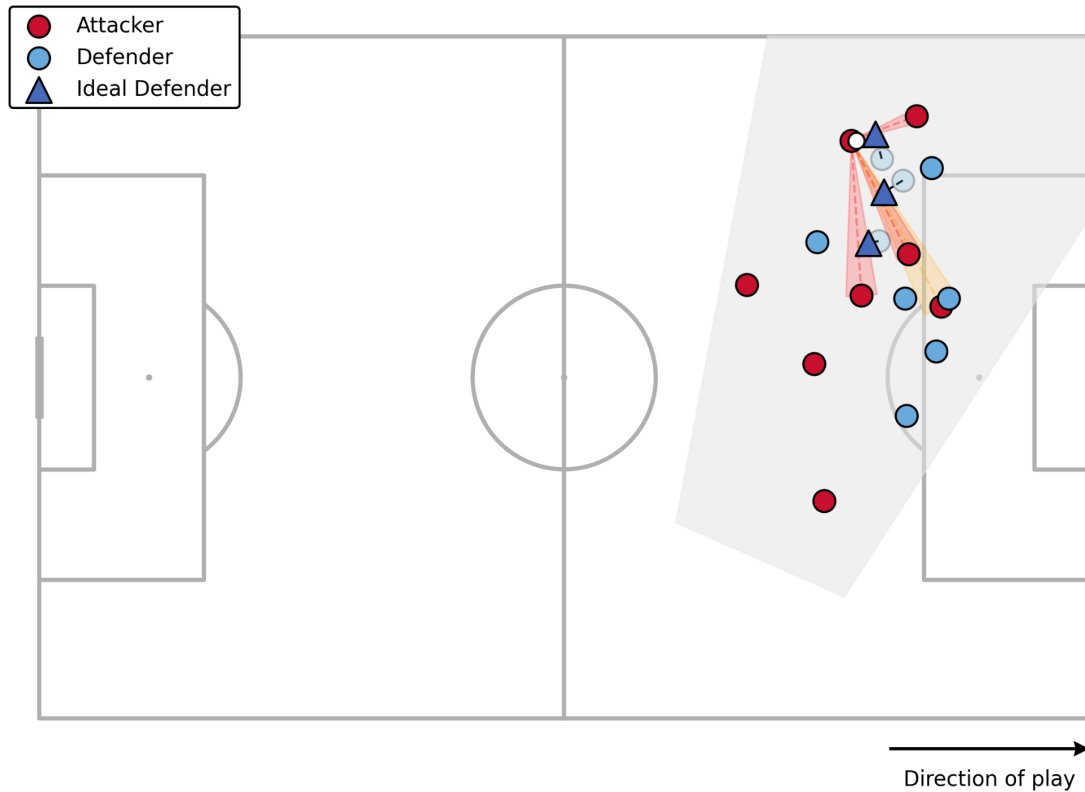


Figure 9: A snapshot illustrating the optimal defensive positioning as result of the linear optimization problem. The positions of three defenders (red dots) are slightly altered (red triangles) to optimally cover the most dangerous attackers (blue dots). The passing cones to the attackers which were already covered in the original scenario (yellow) / are newly covered in the optimal scenario (red) are also shown.

Evaluating Optimization Problem Solutions

We first conduct a basic evaluation of our method by extracting game snapshots preceding the top 1000 most dangerous successful passes from the 2024/25 Serie A season, where the danger consists of the xT value at the receiver’s location. We identify the receiver of the pass, and check if any defender is actively blocking him from receiving the pass (i.e., is located within the cone going from the ball to the pass receiver) in two scenarios: the actual game state, and the alternative game state where defenders are repositioned according to our assignment problem solution. We then count how often this is the case, and report results in Table 4.

Table 4: Comparison of actual vs. suggested defenders positions for the top 1000 most dangerous successful passes of the 2024/25 Serie A season. A pass is blocked if at least one defender is in the cone going from the ball to the actual receiver.

Actual / Suggested	Blocked	Free
Blocked	178	0

Free	789	33
------	-----	----

In 822 out of 1000 passes, no defender was occupying the passing cone of the receiver. Our method suggests a repositioning where the cone is occupied in 789 of these cases. In 33 passes, the receiver comes from behind and is therefore not accounted for in our formulation, which focuses on dangerous attackers between the ball and the goal. In the remaining 178 cases, the receiver was already blocked. Note that it never happens that our formulation “frees up” a dangerous attacker.

Evaluating Cover Shadows Improvements

Finally, we evaluate whether our suggested repositionings reduce the attacking team’s threat, calculated using the full weighted threat formulation from RQ2. Note that differently from RQ2, the optimization problem uses as target defender locations a set of polygons defined by just looking at the xT of each attacker. Hence it does not consider the probability that each attacker receives the ball. However, while we do not explicitly account for pass likelihood or success, our underlying assumption is that the generated solution will inherently discourage passes to high- xT attackers, given that it forces defenders to block them. We now specifically check whether this claim holds as part of our evaluation.

We first analyze the effect of suggested positional adjustments on the threat surface of one specific example, which we show in Figure 10. On the left, we see that in the original game state, the ball carrier has an open corridor towards a teammate closer to the opponents’ goal, who is also the most dangerous attacker according to the threat surface generated by our analysis. As we reposition defenders (on the right), the purple defender is slightly moved to exactly occupy that passing lane. In this alternative scenario, the ball carrier is forced to play a conservative backward pass, which becomes the most likely option. This is reflected in a lower aggregated threat, though in this case the numerical difference is minimal as all players are positioned far from the goal (specifically, it decreases from 0.025 to 0.022).

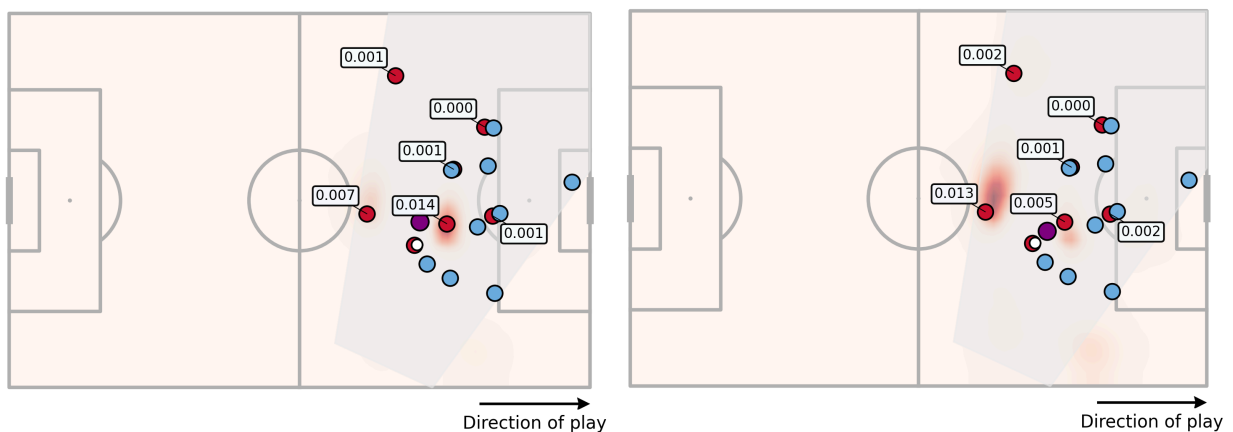


Figure 10: Aggregated threat surfaces for a snapshot before and after applying suggested defender repositionings from the optimization problem solution. Repositioned defenders are purple. Red shading means more threatening pass options. The threat associated with each attacker is the sum of the values of all cells in the Voronoi region of that player.

We broadly evaluate if threat consistently decreases thanks to our repositionings by applying our method on all passes from a full season. We take all moments of interest in the 2024/25 Serie A season (i.e., the interesting snapshots described in our “Data and Definitions” section), and filter out game states where defenders position cannot be improved according to our method, which leaves us with 63,037 snapshots. We then extract updated defenders positions and generate pass selection/success surfaces with SoccerMap, to then compute the aggregated threat for all snapshots where defenders have been repositioned (we compute the threat as explained in RQ2). Finally, we compare the threat after our suggested repositioning with the threat of the original scenario in each of the considered snapshots.

Interestingly, the overall threat is reduced in 75% of the cases, indicating that our repositioning is effective in discouraging dangerous passes. For the remaining cases, we believe modeling our repositioning on basic geometric obstruction of passing cones might be an oversimplification (as players right outside of a cone can still block or influence the pass). Future work should explore how to better model optimal repositioning without relying on simple geometric constraints. A possible line of work involves extending our physics-based approach to define a broader influence zone around the passing lane, capturing all spatial regions where a defender’s presence could realistically deter or intercept a pass, without being constrained to passing cones.

Use Cases

We now use our contributions from RQ1 and RQ2 to study how effective teams and players are at applying cover shadows. Specifically, we present a few insights on a team level by applying our methods on data of the 2024/25 Serie A season. Additionally, we analyze the Belgium-Portugal match from the 2020 European Championship to study performances at a player level.

1. Which teams are better at cover shadowing?

We study teams’ effectiveness at preventing dangerous passes by aggregating their blocking score over all defensive situations from the 2024/25 Serie A season. We consider all game states where the team’s opponents are playing a foot pass between midfield and the defending team box, excluding set pieces and scenarios where no attacker is positioned in front of the ball (as no cover shadow happens there). We then average the total blocking score per match, and show the resulting ranking in Figure 11.

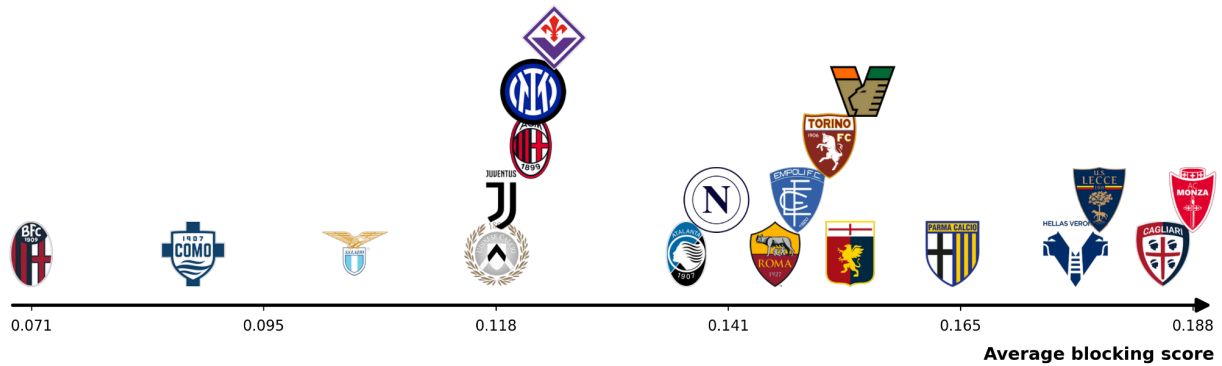


Figure 11: Average blocking score per match for each team of the 2024/25 Serie A season. For each match, we take the sum of the blocking score of all (selected) defensive situations where the team is applying at least a cover shadow.

Notably, teams topping the list finished on the lower end of the table, and were often more focused on limiting their opponents rather than building up valuable plays. Hence one would expect that they pay a lot of attention to cover shadows. On the other end of the spectrum instead are Bologna and Como, two of the best surprises of the year. These were mostly known for an offensive style of play with lots of possession and many players going beyond the ball when attacking. It naturally follows that cover shadows were less of a factor for them.

2. Which players are better at cover shadowing?

We can use our contributions from RQ1 and RQ2 to also quantify a player’s ability to apply effective cover shadows. Together, these metrics capture both how often a player blocks a lane and how much value those blocks contribute defensively. Players who combine a high blocking rate with high threat prevention can be interpreted as especially effective in applying cover shadows.

To do so, we analyze the Belgium vs. Portugal match from the UEFA Euro 2020. We consider freeze frames between midfield and the attacking penalty box in which at least one forward passing option exists. After filtering out frames with incomplete player velocity values, 186 of the 224 snapshots of interest are kept (76 for Belgium and 110 for Portugal in possession), corresponding to 288 potential passes for Belgium and 462 for Portugal.

For each defender, we compute a **blocking rate** as the proportion of frames with a passing lane that is classified as blocked by the defender according to Lane Control:

$$Block\ Rate(d_j) = \frac{\sum_i 1[L(p_a, r_i) \text{ blocked by } d_j]}{Total\ passes\ defended}$$

This simple metric captures how frequently a player obstructs available passing options, independently of whether the opponent actually attempts the pass.

Next, we measure the *impact* of each defender – that is, how much attacking threat is neutralized by each defender’s cover shadow. For each selected snapshot, we estimate the reduction in threat due to the cover shadow applied by each defender (as defined in RQ2) and aggregate these values per player over all snapshots.

Plotting players by their blocking rate (x-axis) and total threat prevented (y-axis) highlights distinct defensive profiles (Figure 12). Players on the lower-right quadrant exhibit the highest block rates, reflecting their frequent role in screening passing lanes. Here we find central midfielders such as Witsel, Tielemans, Palhinha and Moutinho. Players in the higher-left quadrant, instead, block less frequently but prevent a lot of threat. This is the case for fullbacks such as Meunier and T. Hazard, who occasionally find themselves covering crucial defensive areas. Players in the upper-right quadrant both block frequently and prevent high-threat passes, representing the most effective cover-shadowing performers. Here, Sanches represents a mix of the two profiles. Finally, offensive players and central defenders tend to cluster toward the lower-left, indicating limited involvement in preventing dangerous passes.

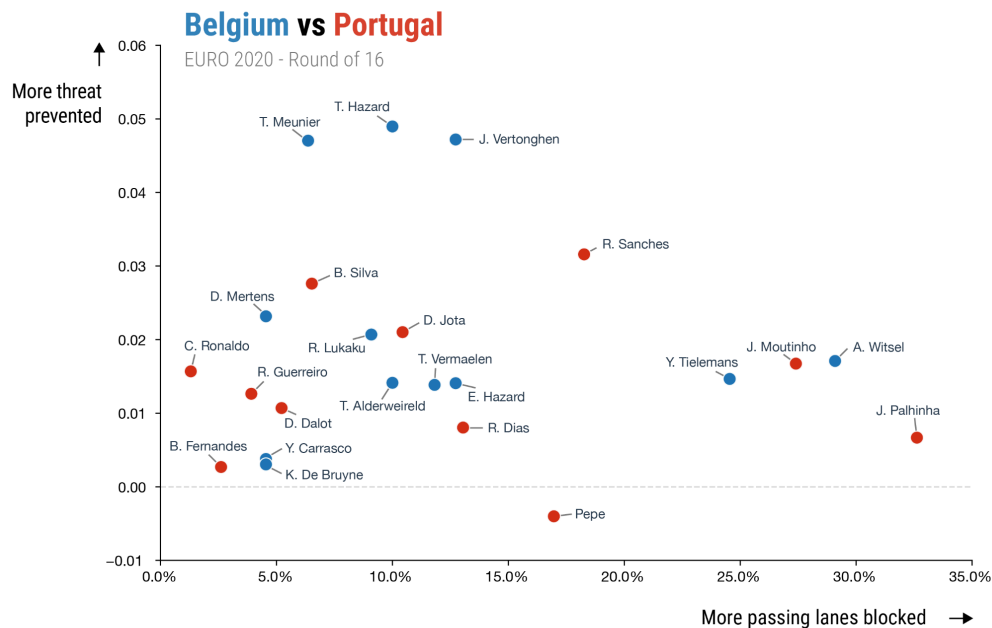


Figure 12: Defensive cover shadowing performances of all players in the Belgium vs. Portugal UEFA Euro 2020 match. Each point represents a player, plotted by their blocking rate (x-axis: proportion of passing lanes blocked) and total threat prevented (y-axis: cumulative xT reduction due to cover shadows).

Related Work

Research in soccer analytics has predominantly focused on offensive actions (i.e., passes, shots, carries). The majority of player valuation models and event-based metrics are designed to assess

how these actions increase the likelihood of scoring [14], [18], [19]. In contrast, defensive contributions have received far less attention. Studies of on-ball actions show a clear bias toward offensive play, leaving defensive involvement underrepresented [20]. This gap is largely due to the nature of defending, which is more about positioning and movement rather than discrete, ball-related events. Defenders exert influence by occupying space, cutting passing lanes, and constraining opponents. Yet, these behaviors were historically difficult to quantify without high-frequency tracking data or freeze-frame snapshots of player positioning.

Therefore, prior work on defensive play analysis using solely event data has mainly focused on notational match statistics of defensive on-ball actions [21], [22] and the valuation of defensive on-ball actions (e.g., interceptions and tackles) by predicting the offensive threat they prevented [32]. However, by solely focusing on the on-ball actions, these approaches only consider a small part of a defender's responsibilities and cannot provide a complete view of their skills.

Using tracking or snapshot data, it becomes possible to evaluate a wider range of positional aspects. To this end, research has mainly looked into general behavioral analysis of players and teams, such as the calculation of centroids and dispersion [23], [24], density [25], and the positioning of defensive lines [26]. In addition, some work has looked into the quantification of *specific* defensive aspects, such as pressure [27], [28], [29], [30] and defensive anticipation [1]. In contrast, the *overall* off-ball defensive influence has been modeled in some frameworks (i.e., expected disruption models) [2], [3] by attributing changes in a team's typical playing style or offensive efficiency to the defensive behavior of opponents. However, these approaches mainly consider man-marking actions and primarily attribute their impact at the team level.

Among works quantifying defensive pressure, J. Bekkers recently proposed a physics-based framework for measuring pressing intensity via time-to-intercept estimation, and outlined a potential extension for passing lanes where pressure is assessed at a point along the line rather than for a moving attacker [11]. Building on this concept, our Lane Control approach parameterizes passing corridors and computes TTI-based blocking probabilities for each point along the corridor.

Most closely related are the recent works of Everett et al. [3], [31]. On the one hand, they provide an approach for calculating optimized defender positions with respect to the threat using a multi-agent Markov Decision Process. However, this uses a coarse-grained grid to represent possible positions, and does not explicitly consider lane-covering defenders nor blockable areas. On the other hand, they analyze a defender's covering abilities by applying graph neural networks (GNNs) and attention mechanisms to measure their influence on the reception probabilities of attackers.

In this work, we build on this line of research and combine both the evaluation and optimization of defensive positioning. Rather than relying solely on latent influence scores, we take a first-principles approach and quantify how defensive positioning, and in particular positional adjustments, change the threat associated with the opponent's possession.

Conclusion

This paper introduced a framework for analyzing cover shadows in soccer. First, we demonstrated the limitations of basic geometric models in detecting blocked passing lanes, and introduced a physics-based model to give a more accurate estimation of *pass blocking probability*. Second, we formulated a novel metric to quantify the threat of the attacking team's passing options, defining *cover shadow effectiveness* as the reduction in this threat. Third, we proposed an algorithm to *optimize defenders' positioning* so that the most dangerous passing lanes are blocked. Therefore, this work provides tools to detect, quantify and optimize the use of cover shadows to prevent dangerous passes.

While the StatsBomb 360 data used in this work provides valuable off-ball information, the provided contextual information is still limited. Future work could extend this framework to full tracking data, which would help in better distinguishing whether a player is a cover shadow defender or a man-marking defender. Full knowledge of players and ball kinematics would also allow for a more accurate physics-based model and more accurate assessments of each passing option's threat. Additionally, having access to player identities would enable a more detailed player-level analysis, such as examining performance across specific regions of the pitch or in different phases of the match. For example, we hypothesize that as fatigue sets in, players become less adept at sustaining optimal positioning. Thus, by tracking a player's effectiveness at maintaining cover shadows, one might be able to identify the onset of fatigue.

Acknowledgements

This research is supported by the Flemish Government under the "Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen" program [AW, LS, MVR, JD], the Research Foundation Flanders (FWO, LC: 11I8125N, MVR: 12A4326N), and KU Leuven Research Fund (iBOF/21/075, C14/24/091 to JD).

References

- [1] Y. Wu and T. Swartz, "Evaluation of off-the-ball actions in soccer," *Stat. Appl. - Ital. J. Appl. Stat.*, no. 2, May 2023, doi: 10.26398/IJAS.0035-008.
- [2] M. Stöckl, T. Seidl, D. Marley, and P. Power, "Making Offensive Play Predictable - Using a Graph Convolutional Network to Understand Defensive Performance in Soccer," *Proc 15th MIT Sloan Sports Anal. Conf.*, 2021, [Online]. Available: <https://www.statsperform.com/wp-content/uploads/2021/04/Making-Offensive-Play-Predictable.pdf>
- [3] G. Everett, R. J. Beal, T. Matthews, T. J. Norman, and S. D. Ramchurn, "Evaluating Defensive Influence in Multi-Agent Systems Using Graph Attention Networks," presented at the 2025 IEEE 12th International Conference on Data Science and Advanced Analytics, 2025.
- [4] M. Van Roy, L. Cascioli, and J. Davis, "ETSY: A Rule-Based Approach to Event and Tracking Data Synchronization," in *Machine Learning and Data Mining for Sports Analytics*, U. Brefeld, J. Davis,

- J. Van Haaren, and A. Zimmermann, Eds., Cham: Springer Nature Switzerland, 2024, pp. 11–23. doi: 10.1007/978-3-031-53833-9_2.
- [5] G. Anzer and P. Bauer, “Expected passes,” *Data Min. Knowl. Discov.*, vol. 36, no. 1, pp. 295–317, Jan. 2022, doi: 10.1007/s10618-021-00810-3.
- [6] P. Robberechts, M. Van Roy, and J. Davis, “un-xPass: Measuring Soccer Player’s Creativity,” in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, in KDD ’23. New York, NY, USA: Association for Computing Machinery, Agosto 2023, pp. 4768–4777. doi: 10.1145/3580305.3599924.
- [7] F. Goes, E. Schwarz, M. Elferink-Gemser, K. Lemmink, and M. Brink, “A risk-reward assessment of passing decisions: comparison between positional roles using tracking data from professional men’s soccer,” *Sci. Med. Footb.*, vol. 6, no. 3, pp. 372–380, July 2022, doi: 10.1080/24733938.2021.1944660.
- [8] D. Vatvani, “Upgrading Expected Goals,” Upgrading Expected Goals. Accessed: Oct. 15, 2025. [Online]. Available: <https://www.hudl.com/blog/upgrading-expected-goals>
- [9] W. Spearman, A. Basye, G. Dick, R. Hotovy, and P. Pop, “Physics-Based Modeling of Pass Probabilities in Soccer,” presented at the MIT Sloan Sports Analytics Conference, Mar. 2017. [Online]. Available: https://www.researchgate.net/publication/315166647_Physics-Based_Modeling_of_Pass_Probabilities_in_Soccer
- [10] W. Spearman, “Beyond Expected Goals,” presented at the MIT Sloan Sports Analytics Conference, Boston, MA, Mar. 2018. [Online]. Available: https://www.researchgate.net/publication/327139841_Beyond_Expected_Goals
- [11] J. Bekkers, “Pressing Intensity: An Intuitive Measure for Pressing in Soccer,” Dec. 30, 2024, *arXiv*: 2501.04712. [Online]. Available: <https://arxiv.org/abs/2501.04712>
- [12] L. Shaw, *Metrica_PitchControl.py*. Github, Friends-of-Tracking-Data-FoTD. Accessed: Oct. 15, 2025. [Online]. Available: https://github.com/Friends-of-Tracking-Data-FoTD/LaurieOnTracking/blob/master/Metrica_PitchControl.py
- [13] F. Haddad, “Building a Pitch Control Model for StatsBomb Event Data,” Medium. Accessed: Oct. 12, 2025. [Online]. Available: <https://medium.com/@fadih3940/building-a-pitch-control-model-for-statsbomb-event-data-e0fbe50cac97>
- [14] K. Singh, “Introducing Expected Threat (xT).” Accessed: Nov. 06, 2025. [Online]. Available: <https://karun.in/blog/expected-threat.html>
- [15] J. Fernández and L. Bornn, “SoccerMap: A Deep Learning Architecture for Visually-Interpretable Analysis in Soccer,” in *Machine Learning and Knowledge Discovery in Databases. Applied Data Science and Demo Track*, Y. Dong, G. Ifrim, D. Mladenić, C. Saunders, and S. Van Hoecke, Eds., Cham: Springer International Publishing, 2021, pp. 491–506. doi: 10.1007/978-3-030-67670-4_30.
- [16] H. W. Kuhn, “The Hungarian method for the assignment problem,” *Nav. Res. Logist. Q.*, vol. 2, no. 1–2, pp. 83–97, 1955, doi: 10.1002/nav.3800020109.
- [17] SciPy Developers, *linear_sum_assignment (Hungarian algorithm)*. (2025). [Online]. Available: https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.linear_sum_assignment.html
- [18] “Introducing On-Ball Value (OBV) • Statsbomb Blog Archive.” Accessed: Nov. 06, 2025. [Online]. Available: <https://blogarchive.statsbomb.com/news/introducing-on-ball-value-obv/>

- [19] T. Decroos, L. Bransen, J. V. Haaren, and J. Davis, "Actions Speak Louder Than Goals: Valuing Player Actions in Soccer," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, July 2019, pp. 1851–1861. doi: 10.1145/3292500.3330758.
- [20] L. Forcher, S. Altmann, L. Forcher, D. Jekauc, and M. Kempe, "The use of player tracking data to analyze defensive play in professional soccer – A scoping review," *Int. J. Sports Sci. Coach.*, vol. 17, no. 6, pp. 1567–1592, Dec. 2022, doi: 10.1177/17479541221075734.
- [21] J. Fernandez-Navarro, C. Ruiz-Ruiz, A. Zubillaga, and L. Fradua, "Tactical Variables Related to Gaining the Ball in Advanced Zones of the Soccer Pitch: Analysis of Differences Among Elite Teams and the Effect of Contextual Variables," *Front. Psychol.*, vol. 10, Jan. 2020, doi: 10.3389/fpsyg.2019.03040.
- [22] H. Lepschy, H. Wäsche, and A. Woll, "Success factors in football: an analysis of the German Bundesliga," *Int. J. Perform. Anal. Sport*, vol. 20, no. 2, pp. 150–164, Mar. 2020, doi: 10.1080/24748668.2020.1726157.
- [23] R. Bartlett, C. Button, M. Robins, A. Dutt-Mazumder, and G. Kennedy, "Analysing Team Coordination Patterns from Player Movement Trajectories in Soccer: Methodological Considerations," *Int. J. Perform. Anal. Sport*, vol. 12, no. 2, pp. 398–424, Aug. 2012, doi: 10.1080/24748668.2012.11868607.
- [24] L. Forcher, L. Forcher, S. Altmann, D. Jekauc, and M. Kempe, "Is a compact organization important for defensive success in elite soccer? – Analysis based on player tracking data," *Int. J. Sports Sci. Coach.*, vol. 19, no. 2, pp. 757–768, Apr. 2024, doi: 10.1177/17479541231172695.
- [25] L. Vilar, D. Araújo, K. Davids, and Y. Bar-Yam, "Science of winning soccer: Emergent pattern-forming dynamics in association football," *J. Syst. Sci. Complex.*, vol. 26, no. 1, pp. 73–84, Feb. 2013, doi: 10.1007/s11424-013-2286-z.
- [26] J. Castellano and D. Casamichana, "What are the differences between first and second divisions of Spanish football teams?," *Int. J. Perform. Anal. Sport*, vol. 15, no. 1, pp. 135–146, Mar. 2015, doi: 10.1080/24748668.2015.11868782.
- [27] S. Merckx, P. Robberechts, Y. Euvrard, and J. Davis, "Measuring the Effectiveness of Pressing in Soccer," *Proc. 8th Workshop Mach. Learn. Data Min. Sports Anal. P 1-15*, 2021.
- [28] I. Bojinov and L. Bornn, "The Pressing Game: Optimal Defensive Disruption in Soccer," *Proc 10th MIT Sloan Sports Anal. Conf.*, 2016.
- [29] G. Andrienko et al., "Visual analysis of pressure in football," *Data Min. Knowl. Discov.*, vol. 31, no. 6, pp. 1793–1839, Nov. 2017, doi: 10.1007/s10618-017-0513-2.
- [30] L. Forcher, L. Forcher, S. Altmann, D. Jekauc, and M. Kempe, "The keys of pressing to gain the ball – Characteristics of defensive pressure in elite soccer using tracking data," *Sci. Med. Footb.*, vol. 8, no. 2, pp. 161–169, Apr. 2024, doi: 10.1080/24733938.2022.2158213.
- [31] G. Everett, R. J. Beal, T. Matthews, T. J. Norman, and S. D. Ramchurn, "Optimising Spatial Teamwork Under Uncertainty," *Proc. AAAI Conf. Artif. Intell.*, vol. 39, no. 22, pp. 23168–23176, Apr. 2025, doi: 10.1609/aaai.v39i22.34482.
- [32] C. Merhej, R. J. Beal, T. Matthews, and S. Ramchurn, "What Happened Next? Using Deep Learning to Value Defensive Actions in Football Event-Data," in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, Virtual Event Singapore: ACM, Aug. 2021, pp. 3394–3403. doi: 10.1145/3447548.3467090.

Appendix

A. Implementation Details for the Gaussian Cone Baseline

We represent each defender d_j by an isotropic [2D Gaussian](#) distribution $N(\mu_j, \sigma_j)$ with mean position $\mu_j = (x_j, y_j)$ and covariance $\Sigma_j = \sigma^2 I$, where σ is the variance multiplier (we use $\sigma = 1$ for 1m^2 variance) and I is the identity matrix. We compute how much the defender's Gaussian distribution covers the cross section of the cone at the defender's projected position on the passing line. For a pass from p_a to r_i of length L meters and unit direction \hat{u} , the parameterized center line is:

$$C(s) = p_a + s\hat{u}, \quad s \in [0, L].$$

For defender d_j , we orthogonally project the defender's mean position μ_j onto the center line to obtain the distance along the pass:

$$s^* = \text{clip}((\mu_j - p_a) \cdot \hat{u}; 0, L).$$

At s^* , the cone has a width $W(s^*)$. The cross section S_j is a line segment of length $W(s^*)$ perpendicular to the pass unit vector \hat{u} and passing through the center line at $C(s^*)$.

Denoting \hat{u}_\perp as the unit vector normal to the pass unit vector \hat{u} at s^* where $(\hat{u}_\perp \cdot \hat{u} = 0)$, the end points of the cross section are:

$$A = C(s^*) - \frac{W(s^*)}{2}\hat{u}_\perp, B = C(s^*) + \frac{W(s^*)}{2}\hat{u}_\perp,$$

The defender's lateral coordinate relative to the segment is

$$u_j = (\mu_j - A) \cdot \hat{u}_\perp, \quad u_j \in [0, W(s^*)]$$

With erf denoting the Gaussian error function, the line integral of the defender's Gaussian over S_j reduces to a 1D normal integral:

$$\frac{1}{2} \left[erf\left(\frac{W(s^*)-u_j}{\sqrt{2}\sigma}\right) - erf\left(\frac{-u_j}{\sqrt{2}\sigma}\right) \right].$$

To ensure comparability across different cross section widths and defender positions, this integral is normalized by its maximum possible value, attained when the defender is standing right in the center of the pass line and $u_j = \frac{W(s^*)}{2}$, giving a denominator of $erf(W(s^*) / 2\sigma)$. The normalized blocking fraction for defender d_j is then:

$$F_\sigma = \frac{erf\left(\frac{W(s^*)-u_j}{\sqrt{2}\sigma}\right) + erf\left(\frac{u_j}{\sqrt{2}\sigma}\right)}{2 erf\left(\frac{W(s^*)}{2\sqrt{2}\sigma}\right)}, F_\sigma \in [0, 1].$$

A pass is considered blocked if any defender's normalized Gaussian blocking fraction exceeds a threshold $\tau = 0.5$:

$$\exists d_j \in D : F_\sigma(p_a, r_i, d_j) \geq \tau.$$

B. Overview of Passing Dataset

We use passes from the Big 5 European leagues in the 2024-2025 season to evaluate our pass block models in RQ1. Specifically, we only retain passes that satisfy the following filters:

- The pass is not a high pass
- The pass originates between midfield and the attacking penalty box
- The pass is progressive (i.e., the x coordinate of the pass end location is ahead of the x coordinate of the pass start location)
- The end location of the pass and the pass receiver are visible in the 360 frame.

The Bundesliga had the highest rate of unsuccessful passes while the EPL had the lowest rate of unsuccessful passes. Note that there are 18 teams and 306 games in both Ligue 1 and the Bundesliga, compared to 20 teams and 380 games in the other three leagues.

Table B1: Distribution of ground truth data from Big 5 European leagues in the 2024-2025 season (total 69,500 passes)

	EPL	Ligue 1	Bundesliga	La Liga	Serie A
Total Passes	16,245	12,815	11,935	14,902	13,603
Successful Passes	16,102	12,665	11,779	14,719	13,431
Unsuccessful Passes	143	150	156	183	172
Actual Pass Block Rate	0.88%	1.17%	1.31%	1.23%	1.26%

C. Evaluating the Pass Selection and Pass Success Models

Evaluating the performance of a pass selection model is not trivial. We follow the procedure from Robberechts et al. [6], extracting the most likely receiver according to the model and comparing it with the actual receiver. We compare with a naive baseline which always predicts the closest teammate as the receiver, an XGB ranking classifier trained on the same handcrafted features from [6], and our SoccerMap model. Table 2 shows that performances are similar between XGBoost and SoccerMap, and coherent with those reported in [6] where SoccerMap is only slightly less accurate than the feature-based model.

Table C1: Average logloss and accuracy of our SoccerMap pass selection model, an XGBoost ranking classifier and a naive baseline for predicting the pass selection probabilities.

	LogLoss	Accuracy
Closest Teammate	-	0.366
XGBoost	-	0.547
SoccerMap	6.521	0.536

We also evaluate our pass success model and report results in Table 3. We again compare with a feature-based XGBoost model trained on the set of handcrafted features from Robberechts et al. Performances are very similar between the two models, and are consistent with the evaluation from [6].

Table C2: Average performance of our SoccerMap model and an XGBoost classifier for predicting the pass success probabilities.

	Precision	Recall	F1	AUC	Brier
XGBoost	0.933	0.938	0.935	0.942	0.073
SoccerMap	0.923	0.942	0.933	0.937	0.076