

Boot It: A Pragmatic Alternative To Build Up Play

Lorenzo Cascioli,¹ Max Goldsmith,² Luca Stradiotti,¹ Maaïke Van Roy,¹ Pieter Robberechts,¹ Maxim Wouters² 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-1020 Brussels, Belgium*

Introduction

A player wins possession deep in their own half, and immediately, shouts echo from the sideline: "BOOT IT!"... Up until the early 2000s, this was a common refrain at soccer matches across all levels and age groups. Back then, the only responsibility of a defender was to ensure goals were not conceded and there was little emphasis on passing skills. Likewise, when a goalkeeper gained control of the ball or took a goal kick, the most frequent expectation was to "send it long".

Today, however, it is increasingly rare to see players kick a long ball to escape pressure and gain a positional advantage on the field at the cost of losing possession. A well-known trend in modern soccer is for teams to patiently build up play from deep within their own half or even from within their defensive box [1]. This strategy can even be favored when facing intense pressure from nearby opponents, where one might expect a more pragmatic approach of clearing the ball to avoid losing possession in a dangerous area.

But is giving up possession always a bad idea? In certain situations, it might be both less dangerous and more effective to try regaining possession close to the opponent's goal and score from there, rather than building up an offensive play from the back. The rise of counter-pressing—where teams immediately apply pressure after losing possession—further highlights the potential benefits of turning over the ball in advanced areas.

In this work, we specifically evaluate the trade-off between trying to retain possession versus intentionally putting the ball out of play near the opponent's goal for a throw-in. While intentionally playing the ball out may seem unconventional, it is not an entirely radical idea. Regularly you see teams employing it as a kick-off strategy.¹ The underlying belief is that an opponent throw-in in the corner of the pitch is worth more than established possession in the center of the pitch.

¹ See for example <https://youtu.be/6svu2FDDbWo?t=710> and https://www.youtube.com/watch?v=5j-lj5_3Cs8&t=720

We see three reasons why kicking the ball out of bounds might be a valuable option, more so than just launching the ball forward. First, professional players should consistently be able to clear the ball out of play near the final third from the midfield area. Second, a throw-in provides the attacking team ample time to regroup and establish an organized press. Third, previous research has shown that possession is more likely to be lost than retained for throw-ins taken within 20 yards (~18m) of a team's own goal [2].

We investigate the trade-off between retaining possession and intentionally putting the ball out using a mix of basic statistical analysis and machine learning. First, we perform a simple analysis that divides the pitch into coarse-grained bins based solely on location (i.e., all other context is ignored). This analysis indicates a slight benefit to the throw-in strategy over retaining possession. Second, we develop a framework based on machine learning to give more nuanced estimates of the payoffs associated with different action choices. By employing machine learning, we can reason about a much richer set of contextual information about the game state such as the locations of players and whether a team is able to apply pressure during the throw-in. Concretely, this enables us to simulate the expected chance of scoring if a team would have kicked the ball out of bounds in certain situations (i.e., that is, we can value a counterfactual action choice). The experimental results of our more nuanced analysis also confirm our original hypothesis that booting a long ball out of bounds is a valid strategy when compared to trying to maintain possession.

Data and Definitions

Our datasets consist of StatsBomb 360 event stream data. This contextualized event stream data is extracted from broadcast video and contains 1) event stream data, and 2) snapshots of player positioning at the moment of each event. The event stream data describes semantic information about the on-the-ball actions, such as which actions are performed, their start and end location, the outcome of the action, which players performed them, and the time in the match they were performed at. To facilitate the analysis, we work with the SPADL representation of this event stream data.² The snapshots include the positions of the players that were visible at the moment of the action, as well as their relationship to the ball carrier (i.e., teammate or opponent).

Our analysis begins by identifying match phases where (i) a player opts to retain possession without a clear opportunity to launch a direct attack and (ii) a team attempts to regain possession on a throw-in. However, distinguishing when a team is deliberately choosing a patient build-up approach is not straightforward. To address this, we use backward passes as a simple indicator. We reason that if a professional soccer player has

² SPADL is a data provider-independent tabular representation for event stream data. It describes each on-the-ball action by the same set of twelve attributes. For more info, see <https://socceraction.readthedocs.io/en/latest/documentation/spadl/spadl.html>.

the option to play a useful progressive pass, they will select it. Hence, if a backward pass was played, it typically means there was no opportunity to launch a direct attack. We identify backward passes by computing the angle between the start and end location of a pass and filtering passes whose angle is $< -5/6 * \pi$ or $> 5/6 * \pi$ (Figure 1).

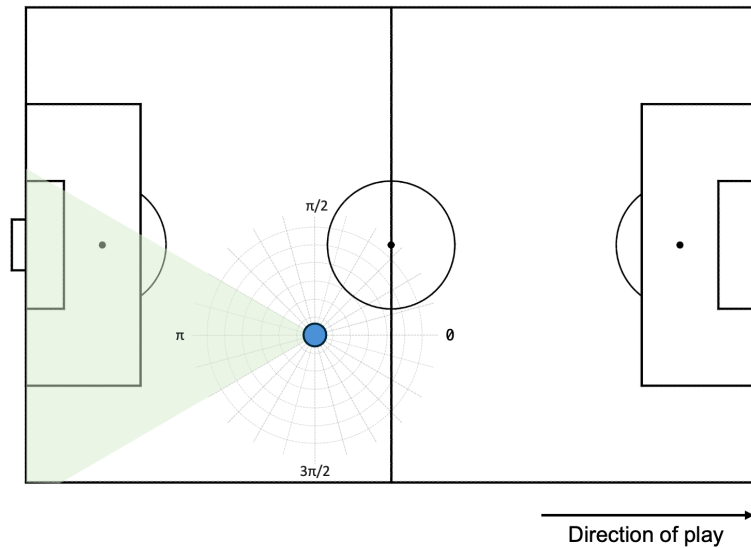


Figure 1. Passes with an angle $< -5/6 * \pi$ or $> 5/6 * \pi$ between the start and end locations are classified as backward passes. Under this definition, any pass originating from the blue dot and ending within the green zone qualifies as a backward pass.

To identify throw-ins where the defending team tries to regain possession, we leverage StatsBomb's provided "*under_pressure*" attribute. This attribute is defined for each on-the-ball action and indicates whether the action was performed while being pressured by an opponent. We classify a throw-in as pressured if (1) the ball is stolen immediately after the throw-in or (2) the team keeps possession, but the first or second pass after the throw-in is labeled "*under_pressure*". An exception is made when the first pass is a long ball, in which case we ignore the second pass.

Our training dataset consists of 1,466,942 passes (of which 351,247 are backward passes) and 63,414 throw-ins (of which 31,300 are pressured). These are extracted from StatsBomb's open source 360 data,³ which include all games from the 2020 and 2024 European Championships, the 2022 World Cup, plus Bayer Leverkusen's title-winning 2023/24 German Bundesliga, together with data from the 2020/21 season of the English Premier League, the Spanish LaLiga and the German Bundesliga, and the 2021/22 and 2022/23 Premier League seasons. A random sample of 20% of passes and throw-ins is used as a validation set when training pass and throw-in value models. In addition, a dataset of 504,962 passes (of which 118,395 are backward passes) and 21,851 throw-ins

³ <https://github.com/statsbomb/open-data>

extracted from the 2022/23 and 2023/24 Italian Serie A seasons has been set aside to evaluate the models and develop the use cases.

Elementary Analysis

We first use a simple binning and counting approach to investigate whether kicking a long ball out of play and applying pressure on the opponent's throw-in leads to a higher probability of scoring compared to making a (possibly risky) backward pass. To simplify our analysis, we assume that kicking the ball out always results in an opponent throw-in.

To compare the two options, we count the number of times a goal is scored within 10 actions [3] of each backward pass or pressured throw-in. In the latter case, we look at goals scored by the team that conceded the throw-in (i.e., the "booting" team recovers the ball and scores from the opponent's throw-in). For a more fine-grained analysis, we divide the passes and throw-ins into five bins according to their end location.

The results presented in Table 1 provide an initial comparison that frames the trade-off of this paper. Namely, the analysis suggests that the probability of scoring from an opponent throw-in in the last bin of the pitch (0.34%) is higher than the probability of scoring as the result of a backward pass in a team's defensive bins (0.18%, 0.20%). Figure 2 shows that a throw-in in the last bin of the pitch (and therefore the long kick that caused it) is more likely to result in a goal compared to a conservative backward pass in the three closest regions to a team's own goal.

Table 1. Probability of scoring in the next 10 actions following a throw-in, or pass out from the back.

Location: X/Touchline Coordinate as the Distance From a Team's Own Goal	From Opponent Throw-in: Probability of Scoring	From Pass: Probability of Scoring
0-21m	0.01%	0.18%
21-42m	0.09%	0.20%
42-63m	0.17%	0.33%
63-84m	0.16%	0.73%
84-105m	0.34%	2.06%

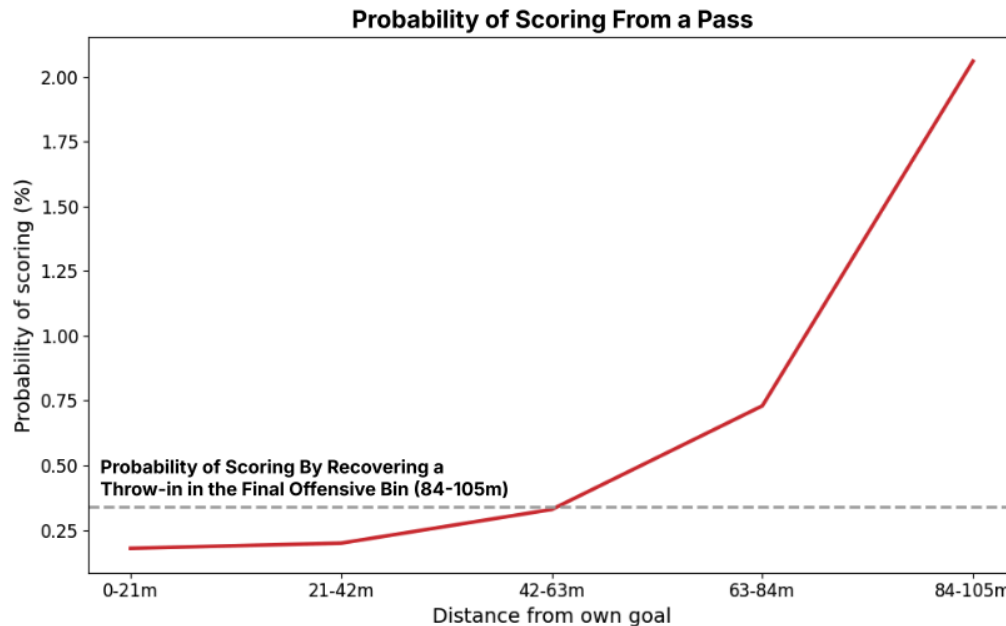


Figure 2. Comparison of scoring probabilities as a result of a pass (red line) or throw-in (dashed line) from different regions of the pitch.

Model-Based Analysis

The elementary analysis suggests that, in certain situations, kicking a long ball out of bounds is a viable alternative to backward passes in areas of the pitch that are either dangerous or do not improve a team’s chances of scoring.

To verify this claim and provide more nuanced estimates of the payoffs associated with different action choices, we employ machine learning techniques to assign a value to each of the two possible choices: 1) a standard backward pass to a teammate, or 2) a long clearance that results in a throw-in for the opponents. Employing machine learning techniques allows us to reason about a much richer set of contextual information about the game states that would result from taking either of these choices and therefore provide more nuanced analyses of the trade-offs.

Model Choice and Settings

Depending on whether a player decides to pass the ball backward or clear it, the resulting game state will be different. The former will result in an open-play possession against a set defense, while the latter will result in a throw-in. Our approach estimates the value of these resulting game states to evaluate the trade-off between both choices.

Estimating the value of a game state is a common problem in soccer analytics. Recent literature [3], [4], [5] has introduced the idea of achieving this by estimating how likely it is for a team to score (i.e., the offensive value) and concede (i.e., the defensive value) a

goal in the near future (i.e., the next few actions) from the current game state. Formally, the value of a game state S_i is then defined as:

$$V(S_i) = P_{score}^k(S_i, t) - P_{concede}^k(S_i, t),$$

where $P_{score}^k(S_i, t)$ and $P_{concede}^k(S_i, t)$ are the probabilities that team t that possesses the ball in state S_i will respectively score or concede in the next k actions. These probabilities can be estimated using a machine-learned model.

A key design choice that differentiates the existing approaches revolves around how to represent a game state. On the one hand, CNN or GNN-based deep learning architectures leverage full spatial tracking data to effectively capture a good representation of the game state [6], [7]. These are applied to prediction tasks such as pass selection, pass success and pass value [6], [8]. On the other hand, we can handcraft features exploiting domain knowledge and use them to train feature-based models. Tree ensemble models such as XGBoost [9] achieve state-of-the-art performance on tabular and event stream data, which typically makes them the chosen option when one can only work with event stream data [3], [4], [8], [10].

While it remains unclear what the best option is for the hybrid 360 data, we follow the work of Robberechts et al. [8], where the authors tried both approaches to build a pass value model and ultimately found that the second option can offer slightly better performance.

Given that goals are rare, multiple works, including the work of Robberechts et al. [8], use xG values as target labels to train possession state value models on a stronger learning signal. We do the same and exploit StatsBomb's xG values for estimating $P_{score}^k(S_i, t)$ and $P_{concede}^k(S_i, t)$: each game state representation is labeled with the combined xG value of all shots in the next k actions taken by the team that respectively possesses the ball and the team that defends in game state S_i . If multiple shots are contained in the action sequence, we extract their combined xG as:

$$xG_{seq} = 1 - \prod_{shot \in seq} 1 - xG_{shot}$$

which evaluates to one minus the probability that no shot in the sequence ends in a goal.

For all XGBoost models, we tune the parameters through a grid search optimizing the max tree depth in {3, 5, 7}, learning rate in {1e-2, 0.1, 0.3}, and number of estimators in {50, 100, 200}. We use early stopping with patience set to 10 boosting rounds.

We use a different set of features to represent the game states that correspond to possession in open play (following a pass) and a throw-in. These feature sets and the resulting models are discussed in the next two sections.

Open Play Game State Value Model

We represent the open play game state following a pass with the best set of features crafted by Robberechts et al. [8] for their pass value model using 360 event stream data. This consists of the traditional set of features used by VAEP [3] with event stream data, which means:

- action type and result
- time of occurrence in the match
- origin and destination location
- the body part used to execute the action
- the team performing the action
- the current goal difference.

We also add features extracted from the 360 frames that capture more context related to the considered pass:

- the number of outplayed players
- the ball interceptability.

The former is computed using a simplified version of Impact's Packing Rate, where a defender is 'packed' if they are located between the ball and the goal before a pass, but further from the goal than the ball after the pass. For the latter, we extract the number of defenders in a 3-meter and 5-meter radius around the pass' start and end location. Note that differently from Robberechts et al. [8], we do not train a model for completed passes and a model for failed passes, as we incorporate the pass result as a feature in our game state representation.

We train these models using all observed passes, and hence do not limit the training phase to backward passes. Thus the models learn to value game states following any kind of pass. We then specifically use them to evaluate game states following backward passes for the purpose of this work.

Table 2 shows the performance of our models. The offensive game state value model's performance is comparable to the performance of the model from Robberechts et al. [8].

Table 2. The performance of our offensive and defensive open play game state value models. Both models are XGBoost ensembles trained on xG values using features from the VAEP framework and additional features extracted from 360 data.

Model	AUC	LogLoss	Brier Score
Open play offensive game state value model	0.790	0.043	0.008
Open play defensive game state value model	0.763	0.010	0.001

Throw-In Game State Value Model

We adopt a similar approach to value game states following throw-ins. We train our models on a set of historical throw-ins, and we use a tailored feature set made of:

- the throw-in’s start and end location
- the current time in the match
- the time elapsed since the previous action (i.e., a proxy for the time to put the ball back in play)
- a Boolean indicator of whether the opponent team is applying pressure on the receiver of the throw-in (see Data section for the definition).

Table 3 reports the performance of the throw-in game state value models.

Table 3. The performance of our offensive and defensive throw-in game state value components. Both models are XGBoost ensembles trained on xG values.

Model	AUC	LogLoss	Brier Score
Offensive throw-in value model	0.727	0.027	0.004
Defensive throw-in value model	0.744	0.008	0.001

Simulating Long Kick + Throw-In as an Alternative to a Pass

The final goal of the model-based approach is to exploit the learned models to answer practical questions around the trade-off between passing the ball backward and kicking it out of bounds in proximity of the opponent’s goal. Given that only one of the two events happens each time a player has the opportunity to choose, our idea is to approach this by starting from an executed backward pass and, together with assigning a value to the game state following the pass, simulating a long ball cleared out of bounds towards the opponent’s goal. We then evaluate the subsequent throw-in game state as the alternative option. Note that while possible in theory, it is hard to make an opposite analysis where

we start from an executed clearance + throw-in scenario and simulate a backward pass, simply because the “boot it!” pattern is rare and harder to pinpoint (often clearances are not a choice, but a necessity).

To complete the analysis, we need a way to simulate a long clearance out of bounds and the subsequent throw-in. Therefore, we start from a snapshot where a player is about to perform a backward pass and we replace the pass by two simulated events: a *clearance* ending out of bounds followed by a *throw-in* for the opponent team.

Long Kick Out of Bounds

We first generate a synthetic *clearance* event starting from the same location as the backward pass. The key point here is estimating how far forward the players can kick the ball. One option is to base this estimate on examples of deliberate long kicks out of bounds in the dataset. However, defining what constitutes a *deliberate* kick out of bounds is challenging. It is difficult to understand from 360 data whether a player’s intention is to simply kick the ball away or to reach a teammate far up the pitch. Therefore, since we are mainly interested in how far players can kick the ball away without aiming for a specific target, we use goal kicks from goalkeepers as a proxy for long kick length. Figure 3 (left) shows the distribution of goal kick lengths using all the 32,392 observed goal kicks in our training set. Specifically, we extract goal kicks longer than 40 meters (assuming shorter kicks are not at full power or aim for a specific teammate), and use their distribution as a proxy for the length of the simulated clearances.

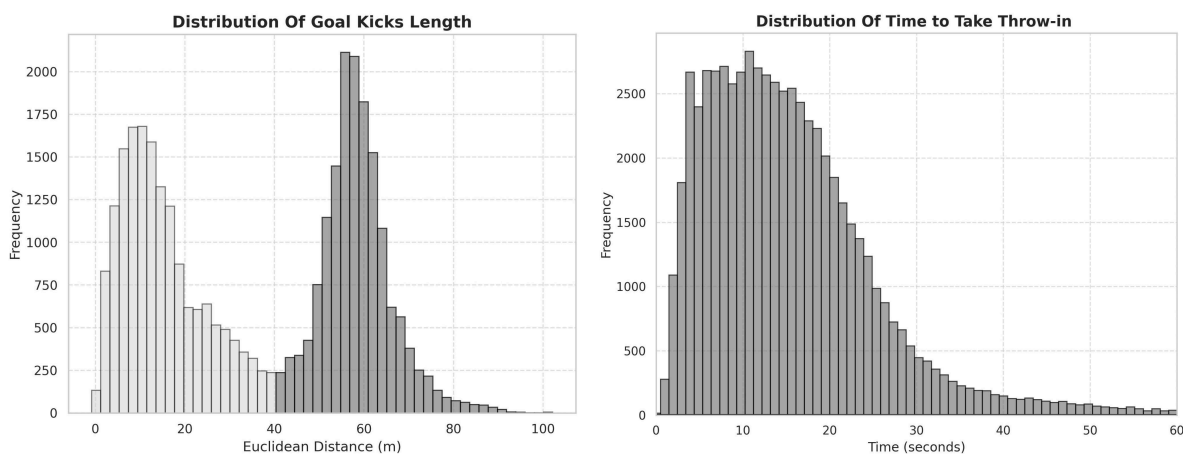


Figure 3. Distributions of goal kicks length (left) and time to take a throw-in (right), extracted using all observed goal kicks and throw-ins from our training set. For goal kicks, we filter out kicks shorter than 40 meters (not at full power) and use the rightmost part of the distribution (which is more darkly shaded) to simulate clearance lengths.

Throw-In

After the clearance has been simulated, we create a synthetic throw-in event at the location where the ball went out of bounds. However, we must account for the time spent between the ball going out of bounds and the throw-in being taken. This is particularly relevant as it is also a feature in the throw-in game state value model. Similarly to what we do for the clearance length, we sample the time needed to take the throw-in from its distribution gathered from the training data (i.e., we extract the time since the previous action from event stream data). The distribution is shown in Figure 3 (right). We do the same for the end location of the throw-in, which is sampled from the set of historical throw-ins close enough to the current one (10 meters tolerance). As an example, Figure 4 shows the distribution of throw-in end locations for a specific throw-in start location.

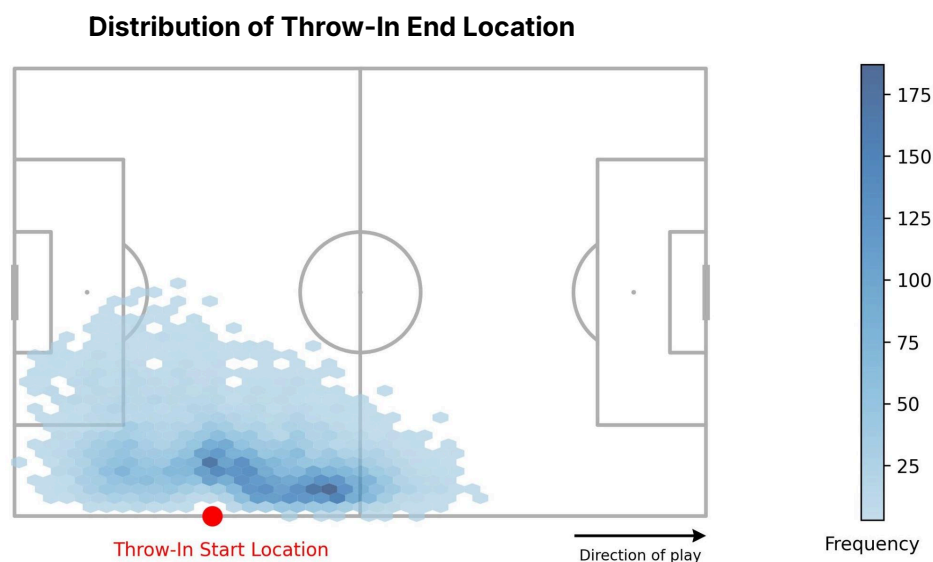


Figure 4. Distribution of throw-in end locations for a specific start location (in red). The distribution includes all observed throw-ins in our training set whose start location was closer than ten meters to the considered start location.

Since we are studying the trade-off between passing the ball and kicking it long to recover it closer to the opponent's goal, in our basic setting we assume that the simulated throw-in is always pressured. However, in the following section, we also experiment with two more scenarios where the throw-in is pressured 50% and 75% of the time. This is to mitigate the (at times unrealistic) assumption that the defending team is always able to apply pressure on the opponent's throw-in after kicking the ball out of bounds.

Finally, note that while we assume that a player will kick a long ball out towards the closest touchline, this simplification does not have a large effect on the analysis. Indeed, while the starting y-location of the subsequent throw-in is included in the throw-in game

state value model, its weight is null as the model learns that it is not a discriminative feature. The only difference is that a clearance of the same length to the further touchline would result in an end location that is slightly further from the opponent's goal.

Use Cases

We use the described models and simulation procedure to compare observed backward passes with simulated clearances+throw-ins for each of the top ten Serie A teams in the 2023/24 season.

Analyzing a Single Match

A first, immediate application of our framework consists of rating all backward passes performed in a match and the corresponding simulated clearances+throw-ins. This allows highlighting a few game states where the choice of passing was not ideal and it would have been preferable to boot the ball. We provide an example in Figure 5. In the presented game state, the team in possession of the ball chose to pass the ball backward. According to our models, however, the team should have avoided the risk of such a pass and should have kicked a long ball out of bounds.

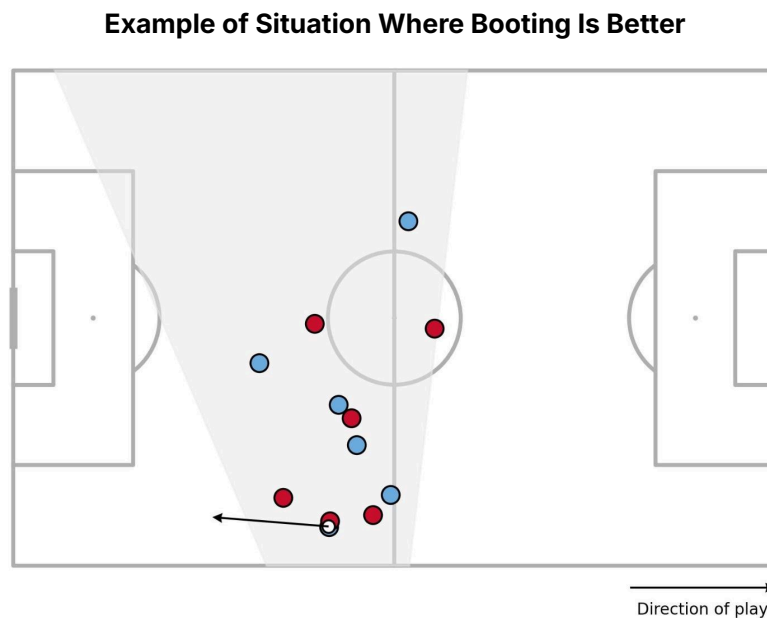


Figure 5. Backward pass extracted from the 2023/24 Inter Milan-AC Milan match. This is the pass in the match where the difference between the two options is maximum, i.e., the value assigned to the simulated clearance+throw-in is higher than the value of the executed pass.

Analyzing a Full Season

Figure 6 shows the distributions of the values attributed to game states following all observed backward passes and throw-ins in our test set. If we exclude passes in a team’s final third, it is rare for a game state to be valued more than 0.01. This is not entirely unexpected: in general, the values are low right after passes and throw-ins because the vast majority of these occur far from the goal, such that the near-term chance of scoring is low.

Values of Game States Following Observed Backward Passes and Throw-Ins

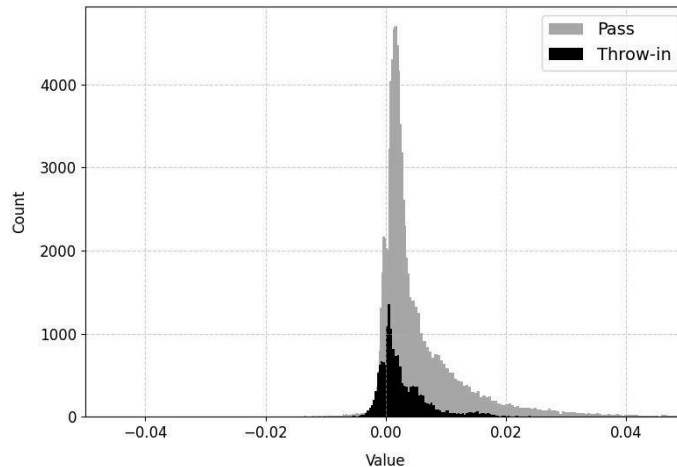


Figure 6. Ratings of backward passes and throw-ins taken in the 2023/24 Serie A season.

Next, we consider all observed backward passes in the defensive third on a team level for the top ten teams in the 2023/24 Serie A season. For each such pass, we simulate booting the ball out of bounds for a throw-in. Figure 7 shows the aggregated difference between the total value of game states following (1) simulated boots and (2) observed passes. In this first analysis, we exclusively focus on passes in a team’s defensive third of the pitch. Intuitively, losing the ball there puts the team in danger as the opponents gain possession close to the goal. At the same time, possessing the ball in the defensive third is not a highly-valued game state per se, as the ball is still far away from the opponent’s goal.

Figure 7 shows that kicking the ball long and out of bounds instead of playing it in the defensive area is indeed advantageous. The trade-off is in favor of the “booting” option for all teams: on average, the difference in possession value over the full season is around +1.08 (i.e., booting results in a bigger reward). Or in other words, this strategy increases a team’s expected goal differential for the season by about 1 goal. If we look at such a difference in terms of expected points, previous work [11] showed that one goal corresponds to approximately one expected point in the league table, which can actually

be a relevant difference impacting (for example) a team's chances to qualify for European competitions or avoid relegation.

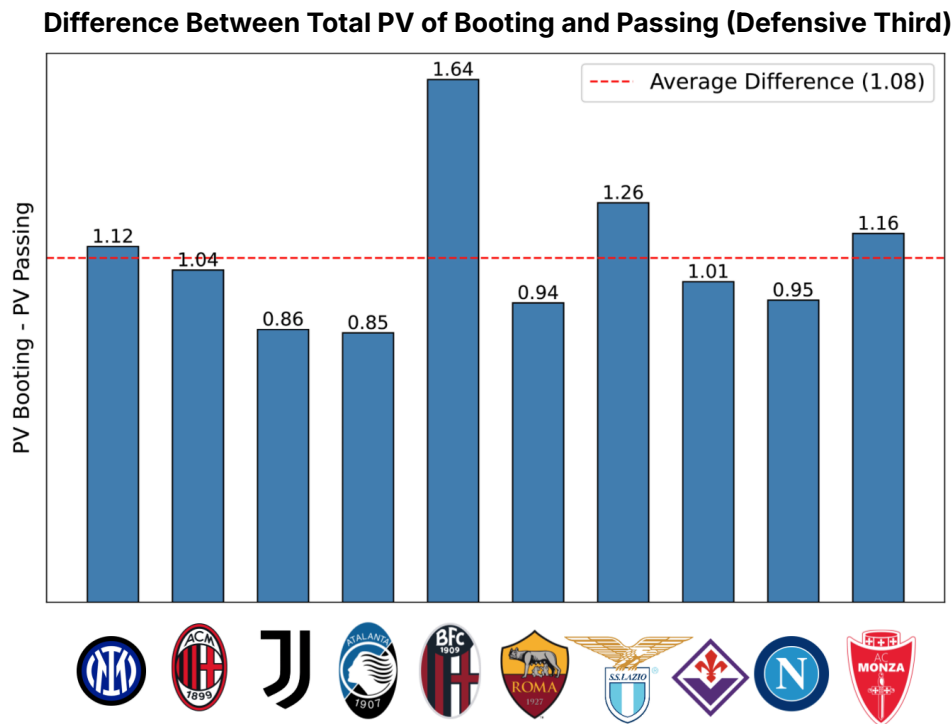


Figure 7. Difference over a season in the total possession value generated by always booting the ball vs always passing it backward, for the top 10 teams and considering all attempted backward passes in a team's defensive third of the 2023/24 Serie A season. A positive value indicates that booting is better.

We then perform the same analysis dividing the full pitch in six bins according to the x-location, and report the average difference between booting and passing for the top ten Serie A teams in each of the bins. Figure 8 shows that as expected, booting the ball is not advantageous when the team has the ball in the offensive midfield: at that point, it is better to retain possession and orchestrate offensive play. However, before the halfway line booting seems to be the preferable option.

Comparing Different Pressing Intensities

In these first experiments, we assume that a team is always able to apply pressure on the opponent's throw-in after kicking the ball out of bounds. This could sometimes be challenging to execute in practice, e.g., due to the opponents putting the ball back in play very quickly. Table 4 presents results for three different fractions of pressured throw-ins. In short, the "100% pressing" column reports the approach from Figure 7 for all three thirds of the pitch. In the other two columns, we see that as the defending team struggles to consistently apply pressure, the "booting" option loses some value. However, in the

defensive third it remains clearly advantageous even if the team is only able to apply pressure in one out of two throw-ins.

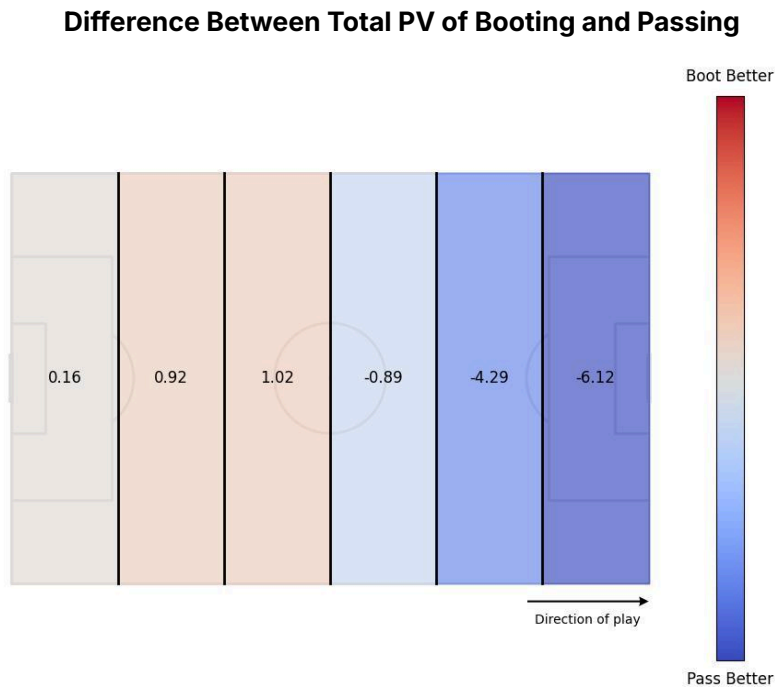


Figure 8. Difference over a season in the total possession value generated by always booting the ball vs always passing it backward. Passes are divided in six bins according to the x-coordinate, and we report the average value for the 10 best teams of the 2023/24 Serie A. A positive value indicates that booting is better.

Table 4. Average difference in total possession value between always booting and always passing for the ten best teams of the 2023/24 Serie A season, for different values of the fraction of pressured throw-ins (proxy for how aggressive the booting team manages to be after clearing the ball).

Area	50% pressing	75% pressing	100% pressing
Defensive third	+ 0.73	+0.91	+1.08
Middle third	-0.90	-0.38	+0.13
Offensive third	-10.88	-10.65	-10.41

Custom Policy

While our analysis indicates that booting the ball is a valid option when a team has the ball in its own half, always choosing the same option is clearly a suboptimal choice. Trade-offs like this are usually better addressed using a mixed policy that selects the option based on the game context (possibly with some randomness for game-theoretic reasons).

We approach this by training an XGBoost policy that predicts which of the two options is best, given a set of historical ratings of passes and throw-ins. Specifically, we:

- Rate all observed backward passes and simulated throw-ins of the 2022/23 Serie A season with the previously trained open play game state value model and throw-in game state value model.
- Use these ratings to train a new XGBoost ensemble to predict which of the two options will receive a higher rating—using the same features as for the open play game state value model.

The ensemble policy achieves 82% accuracy on the test season (2023/24 Serie A). That is, it selects what our models consider to be the best option more than four times out of five.

In Table 5, we compare the performance of the learned policy with respect to our naive simulated boot-only policy and the observed policy (i.e., the team always passes backward), by re-aggregating game state values over the course of the 2023/24 season. In all three zones of the pitch, the learned mixed policy improves over the baselines. In the defensive third, it still seems that always booting the ball is also a valid option. The main difference arises in the middle third of the pitch, where the learned policy is able to consistently improve over the two baselines, generating a possession value difference of around +2 over the course of a season. Instead, as expected, the “booting” option loses value in the offensive third, and the learned policy almost always chooses to pass.

Table 5. Average total possession value for different policies: always passing (what we observe in practice), always booting (our naive simulation), and the learned ensemble policy. Values are aggregated for the ten best teams of the 2023/24 Serie A season.

Area	Observed Policy (Pass-only)	Naive Policy (Boot-only)	Learned Policy
Defensive third	-0.17	+0.91	+1.00
Middle third	+5.77	+5.90	+7.71
Offensive third	+13.18	+2.78	+13.27

Related Work

Throw-ins have been relatively underexplored in soccer analytics but some studies have started to highlight their tactical importance. For instance, McKinley [2] investigated throw-ins to determine their expected value and the likelihood of retaining possession by looking at factors such as the starting and ending location of the throw-in, as well as the time taken to put the ball back into play, and used these findings to offer high-level recommendations for optimizing throw-in tactics. Dona and Swartz [12] investigated the effect of the target location of a thrown-in and found that longer throws confer an advantage in terms of shot creation. More recently, Monte [13] analyzed data from the 2023/24 Premier League season to assess whether certain teams are exploiting throw-ins more effectively to create dangerous situations. Furthermore, Bransen et al. [14] found that teams miss the opportunity to create more danger with throw-ins early in the game. These studies suggest that throw-ins, though not a primary focus in tactical discussions, can still present valuable opportunities for teams to gain an edge.

However, these studies have only looked at throw-ins from an offensive standpoint. No previous work has considered the option of intentionally giving up possession and how to gain an advantage from an opponent throw-in. This setting requires evaluating a trade-off between a sequence of actions.

Analyzing tactical trade-offs in soccer presents significant challenges due to the sport's dynamic, complex, and low-scoring nature. Actions are rarely repeated in the same context and their observed outcomes are often noisy. Researchers have therefore explored various approaches to model and analyze these trade-offs, both at the level of broader tactical strategies (e.g., *direct play* vs *possession play* [15]) and at the level of individual decision-making (e.g., *shoot* vs *pass* [11]).

A first approach involves performing a game-theoretic [16], causal [12], or correlation analysis [15], [17], [18], [19] between the frequency of a given strategy and specific success criteria, such as goals scored. For instance, Charles Reep's early work in soccer analytics [20] found that most goals were scored following fewer than five passes, leading to the (mistaken [21]) conclusion that direct play should be favored over possession-based styles. Although this kind of observational analysis offers some insight into the relative effectiveness of different strategies, it lacks the granularity needed to fully understand the impact of adopting a particular approach and fails to account for contextual factors like the strategy of the opposition.

To get a better insight into the effects of changing a strategy, researchers have leveraged Markov decision processes to build a model of the game [22], [23]. Performance indicators can then be computed via simulation and the model can be tweaked to investigate the effect of tactical changes. While Markov models with a discrete state space offer a useful framework for analyzing player behavior and decision-making, they

struggle to tackle sparse data, meaning that it is hard to incorporate rich contextual information, such as specific game situations or team characteristics.

Lastly, researchers have combined simulation and machine learning to evaluate player decision-making [6], [14], [24], [25]. The approach involves simulating potential future game states one or multiple steps ahead and then assessing the possession value for these simulated states. We adopted this approach since it allows us to reason about a rich set of contextual information about the game state. Moreover, the throw-in is a constrained game state that is relatively straightforward to simulate.

Uniquely, our work compares the effectiveness of different match phases (i.e., open play vs. throw-ins), while previous work has focused on a single, specific match phase.

Conclusions

The presented work argues that the modern emphasis on build-up play might be excessively demonizing long balls out of bounds. Our analysis is based on game state value models and an ad-hoc simulation procedure to compare backward passes with long balls out of bounds. The reported results show that in certain situations, a long kick out of bounds can be a useful move to gain some positional advantage and move the play away from dangerous areas of the pitch. Namely, this appears to be the better alternative in a team's defensive third of the pitch, and a valid option when the ball is in the middle third.

Future work should follow up on this study by addressing a few key limitations. The proposed framework compares observed passes with simulated throw-ins. This is justified by a clear imbalance in the data (there are many more passes than throw-ins), which means that we cannot only work with the observed throw-ins. However, this approach limits the contextual information at disposal of the throw-in game state value model, whose performance could be improved, e.g., using the location of the players at the time of the throw-in. The simulation approach could also be fine-tuned by considering multiple scenarios when simulating a throw-in, as the current procedure simply samples once from the training data distributions.

Acknowledgments

This research is supported by the Flemish Government under the Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen program [LS, MVR, JD] and KU Leuven Research Fund (iBOF/21/075, C14/24/091 to JD).

References

- [1] L. Tharme, "Premier League 2023-24 tactical trends: Return of the No 9, universal pressing and left-footed penalties," *The New York Times*. Accessed: Sep. 25, 2024. [Online]. Available: <https://www.nytimes.com/athletic/5505025/2024/05/21/premier-league-tactical-tren>

- ds-2023-24/
- [2] E. McKinley, "Game of Throw-ins," American Soccer Analysis. Accessed: Sep. 25, 2024. [Online]. Available: <https://www.americansocceranalysis.com/home/2018/11/27/game-of-throw-ins>
 - [3] T. Decroos, L. Bransen, J. Van 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*, Anchorage AK USA: ACM, Jul. 2019, pp. 1851–1861. doi: 10.1145/3292500.3330758.
 - [4] StatsBomb, "Introducing On-Ball Value (OBV)," StatsBomb | Data Champions. Accessed: Sep. 25, 2024. [Online]. Available: <https://statsbomb.com/articles/soccer/introducing-on-ball-value-obv/>
 - [5] J. Muller, "Goals Added: Introducing A New Way To Measure Soccer," American Soccer Analysis. Accessed: Oct. 21, 2024. [Online]. Available: <https://www.americansocceranalysis.com/home/2020/4/22/37ucr0d5urxxtryn2cfhzormdziphq>
 - [6] 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*, vol. 12461, 2021, pp. 491–506. Accessed: Sep. 25, 2024. [Online]. Available: https://link.springer.com/10.1007/978-3-030-67670-4_30
 - [7] M. Stöckl, T. Seidl, D. Marley, and P. Power, "Making Offensive Play Predictable - Using a Graph Convolutional Network to Understand Defensive Performance in Soccer," in *Proceedings of the 15th MIT Sloan Sports Analytics Conference*, 2021.
 - [8] 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*, Long Beach CA USA: ACM, Aug. 2023, pp. 4768–4777. doi: 10.1145/3580305.3599924.
 - [9] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco California USA: ACM, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
 - [10] "Evolving Our Possession Value Framework," Stats Perform. Accessed: Sep. 25, 2024. [Online]. Available: <https://www.statsperform.com/resource/evolving-our-possession-value-framework/>
 - [11] M. Van Roy, P. Robberechts, W.-C. Yang, L. De Raedt, and J. Davis, "Leaving Goals on the Pitch: Evaluating Decision Making in Soccer," in *Proceedings of the 15th MIT Sloan Sports Analytics Conference*, 2021.
 - [12] N. Epasinghege Dona and T. B. Swartz, "Causal Analysis of Tactics in Soccer: The Case of Throw-ins," *IMA J. Manag. Math.*, vol. 35, no. 1, Art. no. 1, Dec. 2023, doi: 10.1093/imaman/dpad022.
 - [13] J. Monte, "Measuring Throw-In Success," StatsBomb | Data Champions. Accessed: Sep. 25, 2024. [Online]. Available: <https://statsbomb.com/articles/soccer/measuring-throw-in-success/>
 - [14] L. Bransen, P. Robberechts, J. Van Haaren, and J. Davis, "Choke or Shine? Quantifying Soccer Players' Abilities to Perform Under Mental Pressure," in *Proceedings of the 13th MIT Sloan Sports Analytics Conference*, 2019.
 - [15] M. Kempe, M. Vogelbein, D. Memmert, and S. Nopp, "Possession vs. Direct Play: Evaluating Tactical Behavior in Elite Soccer," *Int. J. Sports Sci.*, 2014, doi: 10.5923/s.s

ports.201401.05.

- [16] I. Palacios-Huerta, "Professionals Play Minimax," *Rev. Econ. Stud.*, vol. 70, no. 2, Art. no. 2, Apr. 2003, doi: 10.1111/1467-937X.00249.
- [17] P. Power, J. Hobbs, H. Ruiz, X. Wei, and P. Lucey, "Mythbusting Set-Pieces in Soccer," in *Proceedings of the 12th annual MIT Sloan Sports Analytics Conference*, 2018. [Online]. Available: <https://api.semanticscholar.org/CorpusID:189799948>
- [18] B. Noël, J. van der Kamp, and S. Klatt, "The Interplay of Goalkeepers and Penalty Takers Affects Their Chances of Success," *Front. Psychol.*, vol. 12, Mar. 2021, doi: 10.3389/fpsyg.2021.645312.
- [19] L. Shaw and S. Gopaladesikan, "Routine Inspection: A Playbook for Corner Kicks," in *Machine Learning and Data Mining for Sports Analytics*, vol. 1324, 2020, pp. 3–16. Accessed: Oct. 22, 2024. [Online]. Available: https://link.springer.com/10.1007/978-3-030-64912-8_1
- [20] C. Reep and B. Benjamin, "Skill and Chance in Association Football," *J. R. Stat. Soc. Ser. Gen.*, vol. 131, no. 4, p. 581, 1968, doi: 10.2307/2343726.
- [21] M. Hughes and I. Franks, "Analysis of passing sequences, shots and goals in soccer," *J. Sports Sci.*, vol. 23, no. 5, pp. 509–514, May 2005, doi: 10.1080/02640410410001716779.
- [22] N. Sandholtz and L. Bornn, "Markov decision processes with dynamic transition probabilities: An analysis of shooting strategies in basketball," *Ann. Appl. Stat.*, vol. 14, no. 3, Art. no. 3, Sep. 2020, doi: 10.1214/20-AOAS1348.
- [23] M. Van Roy, P. Robberechts, W.-C. Yang, L. De Raedt, and J. Davis, "A Markov Framework for Learning and Reasoning About Strategies in Professional Soccer," *J. Artif. Intell. Res.*, vol. 77, pp. 517–562, Jun. 2023, doi: 10.1613/jair.113934.
- [24] Z. Wang *et al.*, "TacticAI: an AI assistant for football tactics," *Nat. Commun.*, vol. 15, no. 1, Art. no. 1, Mar. 2024, doi: 10.1038/s41467-024-45965-x.
- [25] P. Rahimian, J. Van Haaren, T. Abzhanova, and L. Toka, "Beyond action valuation: A deep reinforcement learning framework for optimizing player decisions in soccer," in *Proceedings of the 16th MIT Sloan Sports Analytics Conference*, 2022.