

exPress: Contextual Valuation of Individual Players Within Pressing Situations in Soccer

Soccer

20251391

Minho Lee¹, Geonhee Jo², Miru Hong³, Pascal Bauer⁴, Sang-Ki Ko⁵

1. Introduction

Pressing is a critical tactical component in modern soccer, enabling teams to disrupt their opponents' build-up play and create advantageous scoring opportunities. Its prominence has grown in recent years, exemplified by high-performing teams like Liverpool and Manchester City, whose aggressive pressing systems have been instrumental in their success [1, 2]. By applying coordinated pressure on the ball carrier and surrounding players, pressing not only forces errors but also facilitates quick transitions to offensive plays.

Consequently, there has been a surge in academic interest in analyzing pressing. Studies have explored various dimensions, including counterpressing [3], pressing actions' effectiveness [4] and impact on game outcomes [5]. Notably, frameworks such as the Valuing Pressure decisions by Estimating Probabilities (VPEP) [5] and subsequent studies [6] have been developed to quantify pressing's value through advanced metrics and machine learning approaches. However, despite these advancements, existing research primarily evaluates pressing outcomes based on event data, often neglecting the nuanced role of player positioning.

The influence of off-ball player positioning during pressing scenarios has only been addressed in a few studies [7]. While some studies assess pressing as a collective effort, only [4] have systematically analyzed whether the spatial arrangement of individual players aligns with optimal tactical expectations using a rule-based quantification of pressure. Moreover, most existing methods lack the granularity needed to determine the appropriateness of player positions during pressing sequences, limiting their applicability for tactical refinement.

This study addresses these gaps by proposing a novel evaluation framework, exPress (Explainable & Expected Press), which centers on player positioning as a key determinant of pressing efficacy. Using existing methodologies, exPress uses XGBoost to model and analyze pressing scenarios. exPress incorporates spatio-temporal data to capture the complex interactions between defenders and attackers. Through interpretable analysis, exPress elucidates how individual player positions and movements influence the outcomes of pressing situations, providing actionable insights into tactical decision-making and player performance evaluation.

¹ Institute for Sports and Preventive Medicine, Saarland University, Email: minho.lee@uni-saarland.de

² Department of Artificial Intelligence, University of Seoul, Email: geonhee@uos.ac.kr

³ Department of Artificial Intelligence, University of Seoul, Email: mirunoyume@uos.ac.kr

⁴ Department of Sports Analytics, Saarland University & Deutscher Fussball-Bund (DFB), Email: pascal.bauer@uni-saarland.de

⁵ Department of Artificial Intelligence, University of Seoul, Email: sangkiko@uos.ac.kr

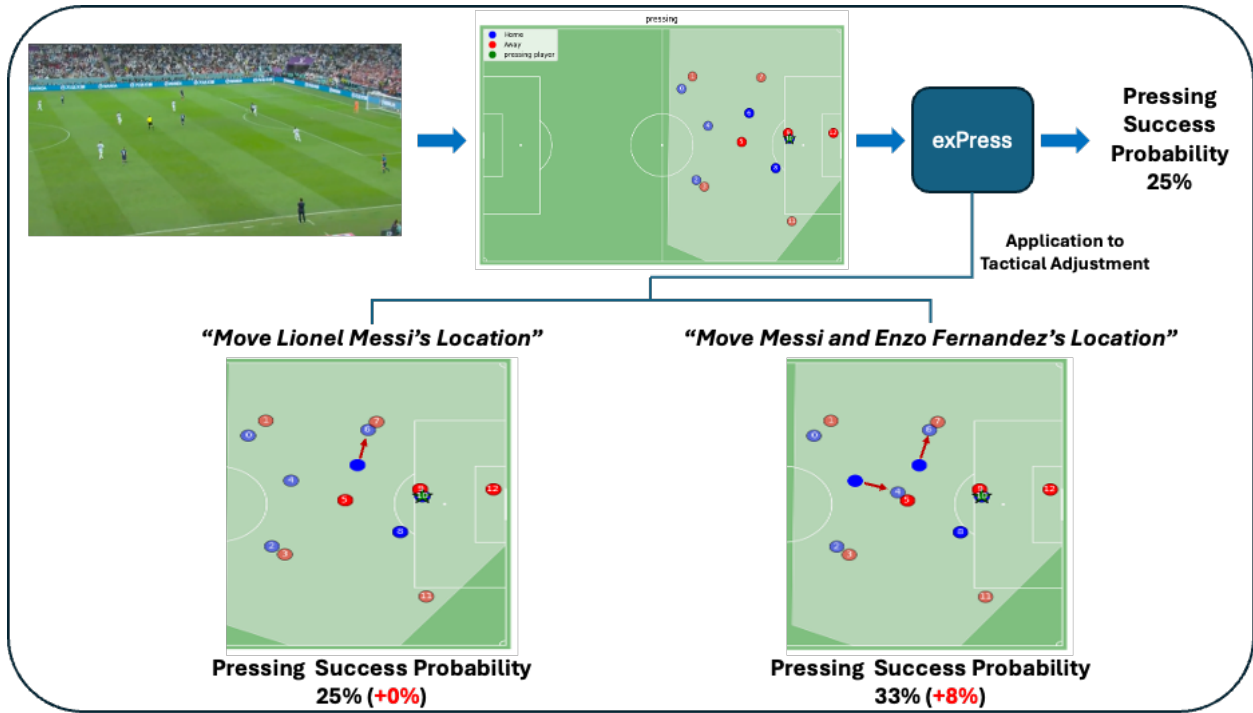


Figure 1. The overall framework of our proposed method. *exPress* evaluates pressing events and player positions by quantifying the effectiveness of pressing based on player positioning and interactions. The framework also provides actionable insights by analyzing counterfactual position adjustments, demonstrating potential increases in pressing success.

2. Related Works

In recent years, advances in data analysis and machine learning technologies have significantly leveraged research on player and team performance in soccer [8, 9]. In literature, positional data predominantly focuses on analyzing offensive movements and strategies, like shooting behavior [10, 11], passing patterns [12, 13, 14, 15] or expected possession value models [16, 17, 18]. More recently, off-ball patterns have been analyzed in the context of set-pieces [19, 20], team formations [21, 22, 23], space control models [24] with various studies also exploring different aspects of pressing [3, 4, 8, 25, 26, 27, 28, 29].

Andrienko et al. [4] introduced a model for quantifying pressure based on distances and angles between players, accompanied by a formula to evaluate pressing intensity. This study also employed visualization techniques to analyze pressing scenarios. Fernandez-Navarro et al. [26] categorized defensive actions into four spatial dimensions and investigated the strategic preferences of teams. Similarly, Low et al. [27] distinguished between high-pressing and deep-defending strategies, utilizing positional data to evaluate metrics such as inter-player distances, team dispersion, and line spacing. Forcher et al. [28] advanced the field by identifying optimal spatial conditions for ball recovery, emphasizing concentrated pressure on ball carriers and proximate areas while intensifying pressure as defensive sequences progressed. Nevertheless, it was limited to analyzing possessions lasting at least five seconds and involving three or more passes, thereby omitting the analysis of counterpressing scenarios, typically occurring within five

seconds of possession loss. Herold et al. [30] quantified and visualized defensive pressure applied to players without ball possession, confirming the hypothesis that reduced pressure on receivers correlates with higher passing success rates. However, the study's focus on specific behavioral concepts, such as Deep Runs and Change of Directions, restricted its ability to examine pressing effectiveness across broader contexts.

More recently, Forcher et al. [31] defined four defensive principles strongly associated with successful ball recovery and explored tactical variations between central and wide areas, highlighting effective defensive structures employed by high-performing teams. Gu et al. [32] introduced a method for quantifying individual player pressure using 3D body motion parameters, creating a "Pressure Matrix" for each player. These matrices were then synthesized into a graph-based Player Pressure Map for evaluating player performance.

The aforementioned studies often lack a comprehensive explanation of the structural dynamics following successful or unsuccessful pressing. To address these limitations, subsequent research has examined the interplay between pressing and subsequent attacking opportunities. Robberechts [5] evaluated pressing decisions through a risk-reward framework and introduced the VPEP framework to contextualize pressing within quantitative analyses. Bauer and Anzer [33] suggested a precise definition of counterpressing, as a team or group tactical behavior aiming to regain the ball immediately after a possession change. They quantified a risk-reward contemplation of counterpressing; however, they were unable to compare its efficiency with alternative tactics, such as fall-ball behavior. Merckx et al. [6] proposed a rule-based method for automatically detecting pressing scenarios. They deconstructed the effects of pressing on ball carriers and their passing options, modeling outcomes such as ball recovery, pass selection, and goal-scoring opportunities, thereby extending the VPEP framework.

Despite these advancements, prior research remains predominantly focused on quantitative performance metrics, often neglecting the influence of player positioning within pressing scenarios. Our approach addresses these gaps by introducing a novel framework for understanding pressing dynamics through an explainable and interpretable approach, bridging theoretical constructs and practical applications.

3. Methods

3.1. Data

Both event and StatsBomb-360 data are provided by Hudl StatsBomb [34]. Event data includes detailed information on in-game actions such as passes, shots, tackles, and pressings—pressings being defined as instances when a player moves to apply pressure on an opposing player in possession of the ball—specifying their time, location, and outcomes. Complementing this, StatsBomb-360 data provides positional information for all players on the field as captured using computer vision, offering detailed insights into player positions visible within the camera's coverage and enabling a more comprehensive analysis of spatial and tactical dynamics.

The dataset comprises 200 matches, covering various competitive levels and contexts. Specifically, it includes 34 matches from the 2023/24 season of Bayer Leverkusen in the 1. Bundesliga, 64 matches from the 2022 FIFA World Cup, 51 matches from the 2024 UEFA European Championship, and 51 matches from UEFA Euro 2020. Including these datasets ensures diverse scenarios for pressing situations, facilitating robust model training and evaluation across different tactical and

competitive environments. To effectively process and analyze the event data from these matches, we utilized the Soccer Player Action Description Language (SPADL) [5] during the preprocessing phase. SPADL is a framework that converts event data provided in different formats by various suppliers, such as Hudl StatsBomb, Opta, and Wyscout, into a unified structure that facilitates analysis. The event data converted through SPADL includes nine attributes: action type, outcome, body part used, start and end locations, start and end times, player identity, and team information. In this study, we combined these nine attributes with the StatsBomb-360 data to precisely quantify pressing situations.

3.2. Implementation Details

We constructed a comprehensive feature set by integrating event data with 360-degree data to effectively represent pressing actions. While event data does only contain information on-ball possessing players, 360 data contain valuable information on a subset of the surrounding players. However, there is one key consideration when using this data source: the number of players captured per event varies, and player identifiers are not provided. For instance, to compute positional and distance information for the three closest players, at least three players from each team must be captured in the data. To handle this, we adopted two approaches: one that directly incorporates missing values—leveraging a tree-based model’s mechanism which automatically learns an optimal default split direction for missing data during tree splitting—and another that imputes missing values with zeros. Further details on the feature construction process are provided in Appendix A.1.

In addition to spatial information, we also accounted for the temporal dynamics between actions. Typically, the action at time a_t precedes the pressing action $p_{t'}$ ($t < t'$); however, there are instances where a_t overlaps with the start of $p_{t'}$. For these cases, we restricted the use of certain features derived from a_t —such as its end position, outcome, and action type to maintain temporal consistency within the feature set.

Building on this process, we define our target to evaluate pressing effectiveness. The primary objective of pressing is generally to recover possession. However, the impact of pressing extends beyond immediate possession recovery. For example, pressing can force the opposing team into disadvantageous positions, thereby increasing their likelihood of losing the ball in subsequent plays, which is also highly valued. Based on this, in this study, we aim to measure the effectiveness of pressing through the probability of regaining possession in the next five seconds, which we denote as xP :

$$xP = P(G = 1 | S_t, p_{t'}) \quad (1)$$

At a given time t , an action in SPADL framework is denoted as a_t . A pressing action occurring at a subsequent time t' is represented separately as $p_{t'}$, where $t < t'$. To capture the contextual information surrounding the pressing action at time t' , we define the game state S_t as a sequence of the previous three actions: $S_t = [a_{t-2}, a_{t-1}, a_t]$.

To assess the effectiveness of a pressing action, we examine whether possession is regained within 5 seconds after t' . A binary outcome variable G is defined, where $G = 1$ (positive label) if possession is regained within 5 seconds, and $G = 0$ (negative label) otherwise. We aim to assess the effectiveness of pressing through the probability of regaining possession.

For modeling pressing outcomes, we employed Extreme Gradient Boosting (XGBoost), a machine learning algorithm that constructs an ensemble of decision trees to model complex relationships. The loss function was tailored to evaluate pressing outcomes effectively, focusing on whether the defending team successfully recovers the ball. Within the dataset, the 2022 FIFA World Cup data was designated as the test set, while the remaining data were used for training. This split resulted in 36,948 training samples and 14,504 test samples, including 10,339 training samples and 4,290 test samples corresponding to instances where ball possession was regained within 5 seconds. Additionally, 20% of the training data was reserved as a validation set to prevent overfitting.

4. Results

We compared the performance of various models to derive pressing value based on player positions and game states. The models included XGBoost, Logistic Regression, Naive Bayes, K-nearest neighbors algorithm (KNN), MLP Classifier, and a CNN-based deep learning approach leveraging the SoccerMap [35] structure. Each model was evaluated based on its ability to predict pressing outcomes, focusing on ball recovery scenarios.

	ROC AUC	Log Loss
Logistic Regression	0.508	0.691
Naive Bayes	0.531	0.809
KNN	0.517	2.919
Random forest	0.593	0.597
MLP Classifier	0.577	0.721
SoccerMap	0.596	0.721
XGBoost	0.607	0.660

Table 1. Results of baseline models in predicting pressing outcomes. This table presents the ROC AUC and Log Loss scores for various machine learning classifiers.

4.1 Comparison to baseline models

Table 1 compares the performance of various machine learning-based classification models. Among these, the best performance was achieved by the XGBoost, which achieved an ROC AUC score of 0.607. The lower performance of the baseline and models compared to previous studies is due to several limitations. Our dataset includes a relatively small number of games, restricting training data diversity. Additionally, positional data is limited to players visible within the camera frame, excluding off-ball movements, and the 360 data lacks information on players not involved in direct actions. These factors reduce contextual completeness, suggesting that future research could improve with larger datasets and more comprehensive player tracking.

4.2 Prediction results for different pressure success conditions

Next, the results of training the XGBoost model with different criteria for successful pressing are compared. When determining the success criteria for pressing, both temporal and spatial factors must be considered. Temporal factors involve evaluating whether ball possession is regained within a certain number of seconds or actions after the pressing begins. Spatial factors involve comparing the distance between the location where the pressing started and the location where possession was regained. For instance, if the point where the pressing began and the point where possession was regained are far apart, it may be difficult to attribute the outcome to the pressing.

	ROC AUC	Log Loss
3 seconds	0.588	0.521
5 seconds	0.618	0.646
7 seconds	0.582	0.661
2 actions	0.712	0.221
4 actions	0.573	0.701
6 actions	0.586	0.669
5 seconds in 3 meters radius	0.618	0.213
5 seconds in 5 meters radius	0.618	0.316
5 seconds in 9 meters radius	0.621	0.431

Table 2. Performance comparison based on different pressing success criteria. This table compares the predictive performance of the XGBoost model when different success criteria for pressing are applied. The evaluation includes temporal thresholds as well as spatial constraints. The ROC AUC and Log Loss metrics illustrate how the incorporation of both temporal and spatial factors affects the model's assessment of pressing effectiveness.

5. Application

5.1 Providing positional feedback for better pressing or press-resistance

In the following, we demonstrate that our proposed model, exPress, can analyze pressing scenarios by evaluating the impact of player positioning on the probability of successful pressing actions. By simulating positional adjustments, the model provides insights into how player movements optimize pressing efficiency and limit the passing options available to opponents. Iteratively modifying player locations allowed us to observe changes in pressing success probabilities, thereby identifying which positional improvements enhance defensive effectiveness and which configurations leave opponents with viable passing alternatives. In this analysis, the simulations of player positioning were informed by video data, through which the authors discussed and determined the positions that would have led to successful pressing or allowed the pressing to be neutralized.

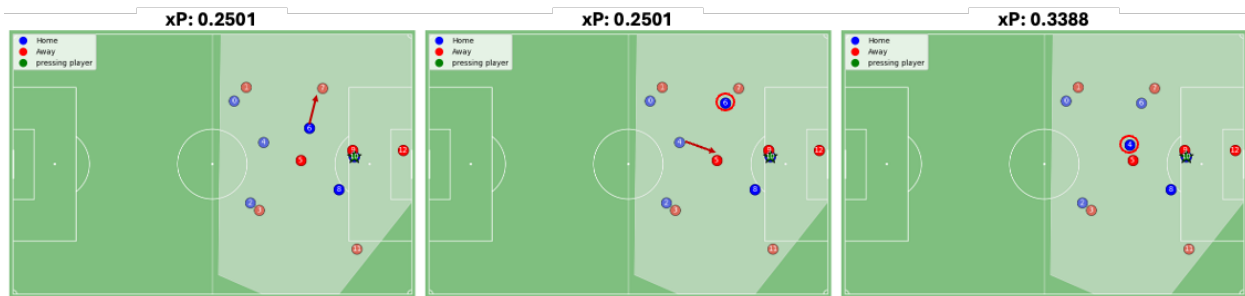


Figure 2. Application example of positional adjustment on a pressing scenario. The first image (left) depicts the actual pressing scenario, and the second image shows an adjustment, moving player 6 closer to player 7. Lastly, the third image (right) demonstrates a further adjustment with player 4 marking player 5, which increases the pressing success probability to 0.3388.

Figure 2 illustrates the progression of a pressing scenario and the effectiveness of positional adjustments as evaluated by the proposed pressing model, exPress. In the first image (left), the actual pressing scenario is depicted, where Home team (blue) player 10 is pressing Away team (red) player 9. However, Home team player 6 is positioned too far from Away players 5 and 7, leaving both players as viable passing options. As a result, the pressing score (xP : 0.2501) is low, reflecting the inefficiency of the pressing action. This inadequacy is further highlighted by the outcome of the play, where Home's pressing fails, and Away successfully takes a shot. To adjust this situation, in the second image (center), Home team player 6's position is adjusted to be closer to Away player 7. Despite this change, the predicted xP remains unchanged at 0.2501. This can be attributed to Away player 5 being left entirely free, and the positional adjustment inadvertently increases the freedom of player 5 to receive a pass, thereby negating any potential defensive improvement. Therefore, as in the third image (right), a further adjustment is made by repositioning Home team player 4 closer to Away player 5. This adjustment significantly increases the predicted pressing score (xP : 0.3388). By marking Away player 5, the pressing pressure is distributed more effectively, restricting the opponent's passing options and increasing the likelihood of a successful pressing outcome.

We also demonstrate how the proposed model can also be utilized to improve press-resistance for teams under pressure, as shown in Figure 3. In the first image (left), Home team (blue) player 10 effectively presses Away team player 12, resulting in a high pressing success probability of 0.4744. In this situation, the Away team faces a significant challenge in breaking the press due to limited passing options and constrained movement. To explore a potential press-resistance strategy, Away team player 5's position was adjusted to move toward the left side of the pitch, creating an additional passing path and offering a potential outlet to relieve the pressure. This positional adjustment is depicted in the second image (right), where the Away team's player 5 provides a more viable passing option to escape the press. Following this adjustment, the predicted pressing success probability drops to 0.4675, indicating a reduced likelihood of the Home team successfully executing their pressing action.

Taken together, exPress optimizes both pressing effectiveness and press-resistance by simulating positional adjustments and their impact on play dynamics. By providing actionable tactical insights, it enables teams to enhance defensive pressure or strategically evade it, supporting real-time decision-making in pressing scenarios.

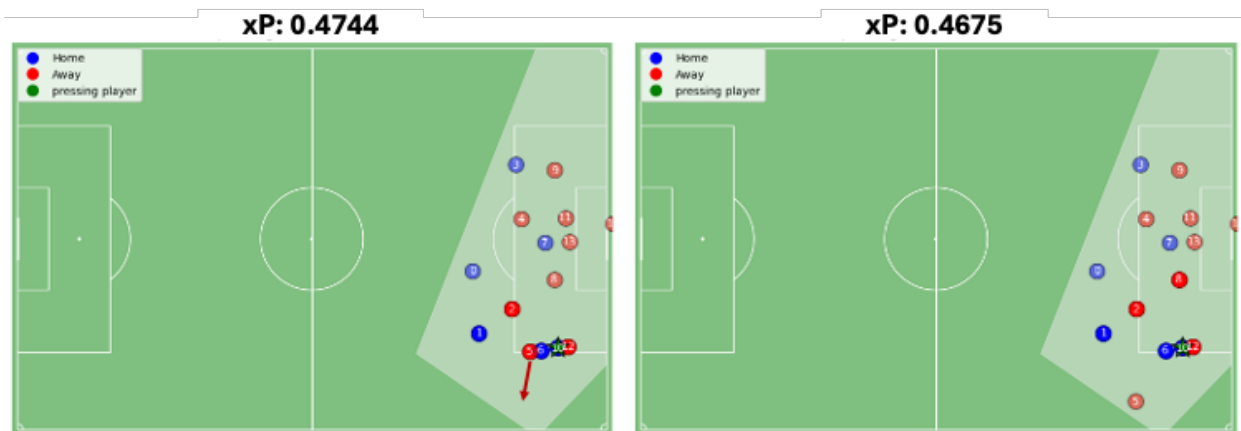


Figure 3. Application example of positional adjustment on pressing success. The first image (left) depicts the actual pressing scenario, where Home team (blue) player 3's position allows a passing option to Away (red) team player 5, resulting in a pressing success probability of 0.4744. The second image shows a positional adjustment, placing player 3 closer to Away player 5 and between players 1 and 5, effectively restricting passing options and xP decreased (0.4675).

5.2 Evaluation of pressing effectiveness of teams

We analyzed pressing performance at the 2022 Qatar World Cup using the proposed model, focusing on teams' ability to regain possession under pressure. To evaluate team-level pressing performance, we aggregated xP values across all pressing situations for each team. Pressing efficiency was assessed using xP and xP Difference, which measure a team's deviation from expected outcomes, providing insights into their overall pressing effectiveness. The xP Difference is calculated as the difference between a team's actual pressing success result and their expected pressing success rate (xP). A positive xP Difference indicates that a team successfully regained possession in situations where pressing was statistically less likely to succeed. In contrast, a negative xP Difference suggests ineffective pressing execution, where teams struggled to capitalize on favorable pressing opportunities.

Figure 4 visualizes pressing efficiency of teams at the 2022 World Cup. Teams located in the top-right quadrant (e.g., Uruguay, Netherlands) demonstrated both high xP per 90 and xP Difference per 90, indicating that their pressing actions were both frequent and effective. In contrast, Japan and Iran recorded high xP but a negative xP Difference, suggesting that while they engaged in pressing frequently, their execution was not as effective as expected. This discrepancy may reflect structural weaknesses in their pressing schemes or an inability to convert pressure into possession recovery.

It is important to note that, given the unique structure of the World Cup, teams face opponents of varying quality, which can influence xP values. The strength and style of the opposition may impact the effectiveness of pressing actions, meaning that teams playing against higher-caliber opponents might naturally exhibit different xP metrics compared to those facing less formidable challengers. Therefore, when interpreting these results, the level of the opposing team should be considered as it may skew the pressing efficiency metrics.

How Teams Pressed in 2022 World Cup

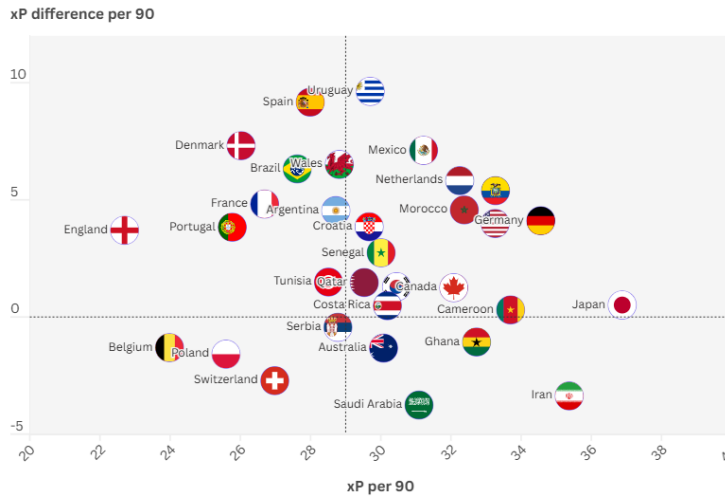


Figure 4. An illustration of xP and xP difference for the countries that participated in the 2022 Qatar World Cup. This visualization quantifies how effectively each team executed their pressing actions and how their performance aligned with the expected outcomes.

5.3 Pressing / Press-resistance ability of individual players

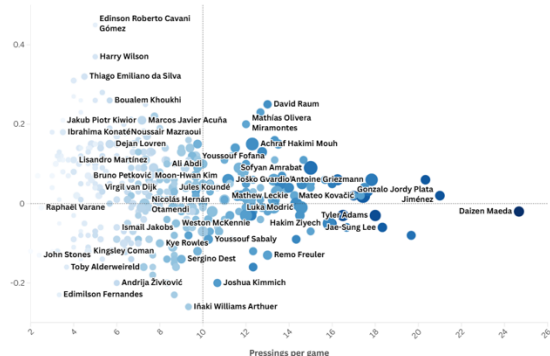
Pressing effectiveness and press resistance are fundamental aspects of modern football. A player with strong pressing ability can force turnovers, disrupt the opponent's build-up play, and create scoring opportunities. Likewise, a player with strong press resistance can maintain possession under pressure, exploit defensive gaps, and facilitate counterattacks. To analyze these dynamics, we applied our model to player data from the 2022 Qatar World Cup, and quantified both pressing and press resistance abilities.

Pressing Effectiveness

Players with 3+ Appearances in the 2022 World Cup

xP_per_game 0.63 5.57

xP difference per pressing



Pressing Effectiveness

Players with 3+ Appearances in the 2022 World Cup

xP 2.04 33.36

xP diff.

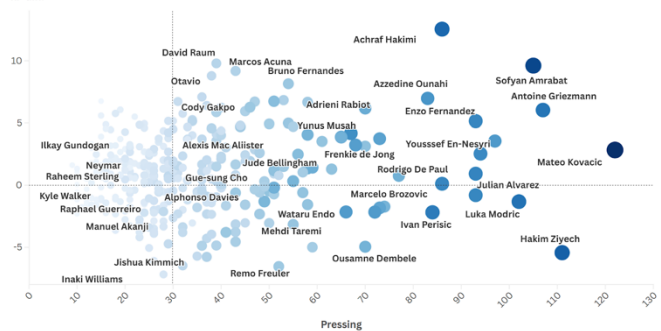


Figure 5 and 6. An illustration of pressings per game and xP difference per pressing for the player who played 3 or more games in the 2022 Qatar World Cup. Figure 5 (left) Displays pressings per game and xP Difference per pressing, Figure 6 (right): Uses cumulative pressings and cumulative xP Difference, the size and color intensity of each point further underline the xP.

Figure 5 (left) illustrates pressing frequency per game and xP Difference per pressing action. Players in the top-right quadrant (e.g., David Raum, Enzo Fernandez) pressed frequently and effectively, recovering possession in key moments. Sofyan Amrabat, a key midfielder for Morocco, consistently applied defensive pressure with high success rates. In contrast, Daizen Maeda and Idrissa Gana Gueye (bottom-right quadrant) frequently engaged in pressing but with lower efficiency, suggesting that many of their attempts occurred in suboptimal situations.

Figure 6 (right) provides insights into cumulative pressing actions, highlighting differences in player roles. Achraf Hakimi and Hakim Ziyech, both key players for Morocco, demonstrated contrasting pressing effectiveness. Hakimi's pressing actions consistently exceeded expectations, whereas Ziyech, despite pressing frequently, showed lower xP Difference, indicating that his pressing attempts were less impactful. This distinction underscores how pressing volume and efficiency contribute differently to a team's defensive strategy. The xP Difference metric thus provides a deeper understanding of individual pressing effectiveness, distinguishing between high impact pressing and mere activity.

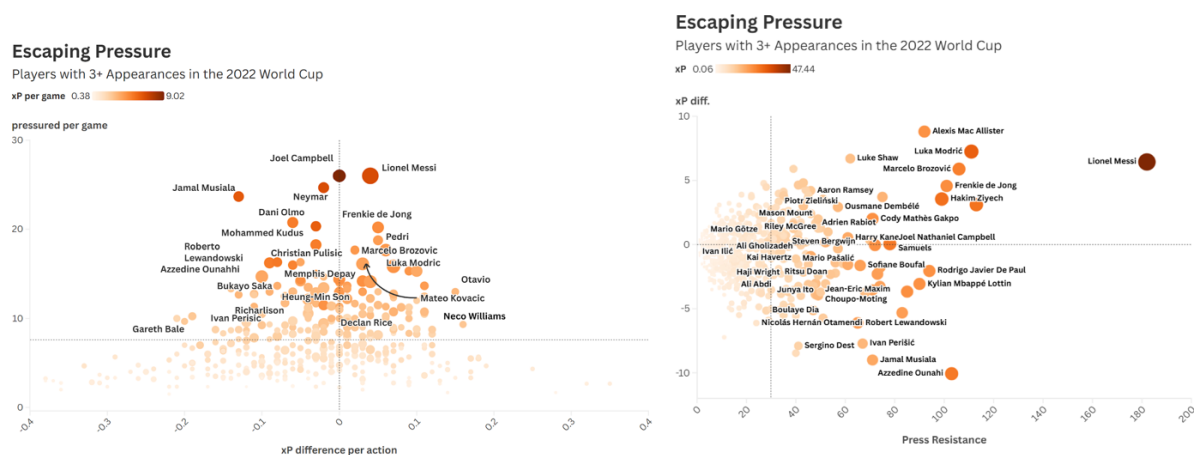


Figure 7 and 8. An illustration of press-resistance ability among players with at least three appearances in the 2022 Qatar World Cup. Figure 7 (left) displays pressured actions per game and xP Difference per pressured action, while Figure 8 (right) shows cumulative pressured actions and cumulative xP Difference. The size and color intensity of each point represent xP , highlighting variations in press resistance across players.

Players frequently targeted by pressing often play pivotal roles in build-up play. However, their ability to resist pressure varies. In Figure 7, Lionel Messi, positioned in the top-right quadrant was frequently pressed yet exhibited exceptional efficiency in evading pressure, significantly outperforming model expectations. His ability to retain possession under pressure made him a crucial playmaker for Argentina. In contrast, Jamal Musiala, also frequently targeted by pressing (top-left quadrant), struggled to navigate such situations as effectively. His lower xP Difference suggests that he was often forced into less favorable decisions or turnovers under pressure. A similar pattern is observed in Figure 8. Despite facing a comparable number of pressing actions, central midfielder Alexis Mac Allister and Azzedine Ounahi displayed contrasting abilities to withstand pressure. Mac Allister navigated pressing situations with composure, frequently progressing play. Whereas Ounahi found himself struggling under defensive pressure, with a lower xP Difference.

These findings underscore the varied abilities of players to handle pressing and highlight potential tactical adjustments for optimizing pressing efficiency and press resistance. By quantifying these aspects, our model offers valuable insights for refining team strategies and individual player roles in high-pressure situations.

6. Discussion

While our model effectively quantifies pressing and press-resistance abilities, certain limitations must be acknowledged. One of the limitations of this approach is that pressing efficiency heavily depends on the opponent's team strength. When facing stronger teams, higher levels of defensive pressure are required to regain possession compared to weaker teams. This means that xP Difference may not solely reflect a team's pressing quality but also the difficulty posed by their opponents (e.g., individual player ability and tactical adaptability). Furthermore, the tournament format limits the number of possible team matchups, meaning that some teams may face disproportionately strong or weak opponents. As a result, pressing efficiency measurements may not fully generalize across different levels of competition.

Another limitation is that player identities are not explicitly considered in the model. Currently, player positions (x/y coordinates) are used as inputs without incorporating individual player characteristics such as playing style, experience, or tactical tendencies. As a result, pressing and press-resistance evaluations rely solely on spatial positioning rather than personalized player attributes. Additionally, the way x/y coordinates are fed into the model could introduce potential biases. While our approach incorporates various features beyond raw coordinates—such as distances to the nearest opponent and teammate, angles relative to the goal, and other contextual information—it does not fully capture player positioning in relation to the ball in possession.

Despite integrating multiple spatial features, our model does not explicitly structure inputs in a way that directly relates player positioning to the ball. Future improvements could refine this by restructuring input data to incorporate formation-based positional adjustments or considering the spatial distribution of opposing players within potential passing lanes.

Furthermore, the nature of StatsBomb-360 data presents inherent limitations due to missing player information. Since the dataset is derived from broadcast camera angles, not all players are always visible. As a result, some pressing sequences may be incomplete, potentially affecting xP calculations. Addressing this issue would require additional data sources, such as tracking data, to ensure a more comprehensive representation of all players involved in pressing sequences.

7. Conclusion

In this study, we introduce a novel framework, exPress, designed to evaluate pressing scenarios in soccer by analyzing spatial interactions between the player in possession and the surrounding defender. The model leverages XGBoost to provide an explainable and interpretable approach for quantifying pressing success probabilities (xP). Through experimentation, we demonstrated that exPress effectively identifies optimal player positions to enhance pressing outcomes while also suggesting positional adjustments to improve press-resistance for teams under pressure. The application of exPress revealed its capability to simulate and evaluate positional adjustments, providing actionable insights for both pressing and resisting pressure. Our results highlight the

practical utility of the framework in informing tactical decisions, optimizing team strategies, and bridging the gap between academic research and practical soccer analytics. Despite its promising contributions, the model's performance was influenced by limitations in data volume and scope, as it utilized a relatively small dataset and positional information restricted to players visible within the camera frame. Future research could address these challenges by incorporating larger datasets and more comprehensive player tracking data, further enhancing the model's accuracy and applicability.

Acknowledgments

This research was conducted as part of a project supported by the Korea AI Research Society for Sports (KAISports). We gratefully acknowledge Hyunsung Kim (KAIST & Fittogether) and Jaemin Lee (Samsung Electronics) for their significant contributions to the conceptualization of this study.

References

- [1] Sam Lee. "Manchester City's Perfect Goal Under Pep Guardiola." The Athletic. April 5, 2024
- [2] UEFA.com "Champions League performance insights: The power of Liverpool's press" UEFA.com. September 20, 2024
- [3] Pascal Bauer. "Automated Detection of Complex Tactical Patterns in Football." (2021).
- [4] Gennady Andrienko et al. "Visual analysis of pressure in football." Data Mining and Knowledge Discovery. 31(6), 1793–1839, 2017
- [5] Pieter Robberechts. "Valuing the art of pressing." StatsBomb Innovation in Football Conference. Vol. 11. 2019.
- [6] Simon Merckx et al. "Measuring the effectiveness of pressing in soccer." In Machine Learning and Data Mining for Sports Analytics: 8th International Workshop, MLSA 2021.
- [7] Forcher, Leander, et al. "The use of player tracking data to analyze defensive play in professional soccer-A scoping review." International Journal of Sports Science & Coaching 17.6 (2022): 1567-1592.
- [8] Herold et al. "Machine learning in men's professional football: Current applications and future directions for improving attacking play." International Journal of Sports Science & Coaching 14.6 (2019): 798-817.
- [9] Goes, F. R., et al. "Unlocking the potential of big data to support tactical performance analysis in professional soccer: A systematic review." European Journal of Sport Science 21.4 (2021): 481-496.
- [10] Anzer, Gabriel, and Pascal Bauer. "A goal scoring probability model for shots based on synchronized positional and event data in football (soccer)." Frontiers in sports and active living 3 (2021): 624475.
- [11] Anzer, Gabriel, Pascal Bauer, and Ulf Brefeld. "The origins of goals in the German Bundesliga." Journal of Sports Sciences 39.22 (2021): 2525-2544.
- [12] Anzer, Gabriel, and Pascal Bauer. "Expected passes: determining the difficulty of a pass in football (soccer) using spatio-temporal data." Data mining and knowledge discovery 36.1 (2022): 295-317.
- [13] Chawla, Sanjay, et al. "Classification of passes in football matches using spatiotemporal data." ACM Transactions on Spatial Algorithms and Systems (TSAS) 3.2 (2017): 1-30.
- [14] Power, Paul, et al. "Not all passes are created equal: Objectively measuring the risk and reward of passes in soccer from tracking data." Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining. 2017.
- [15] Spearman, William, et al. "Physics-based modeling of pass probabilities in soccer." Proceeding of the 11th MIT Sloan Sports Analytics Conference. Vol. 1. 2017.
- [16] Fernandez, Javier, and Luke Bornn. "Wide Open Spaces: A statistical technique for measuring space creation in professional soccer." 12th MIT Sloan Sports Analytics Conference. Vol. 2018. 2018.
- [17] Fernández, Javier, Luke Bornn, and Dan Cervone. "Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer." 13th MIT Sloan Sports Analytics Conference. Vol. 2. 2019.
- [18] Spearman, William. "Beyond expected goals." Proceedings of the 12th MIT Sloan Sports Analytics Conference. 2018.
- [19] Shaw, Laurie, and Sudarshan Gopaladesikan. "Routine inspection: A playbook for corner kicks." Proceedings of the 15th MIT Sloan Sports Analytics Conference. 2021.
- [20] Bauer, Pascal, Gabriel Anzer, and Joshua Wyatt Smith. "Individual role classification for players defending corners in football (soccer) Categorisation of the defensive role for each player in a corner kick using positional data." Journal of Quantitative Analysis in Sports 18.2 (2022): 147-160.

- [21] Bauer, Pascal, Gabriel Anzer, and Laurie Shaw. "Putting team formations in association football into context." *Journal of Sports Analytics* 9.1 (2023): 39-59.
- [22] Forcher, Leon, et al. "Shedding some light on in-game formation changes in the German Bundesliga: Frequency, contextual factors, and differences between offensive and defensive formations." *International Journal of Sports Science & Coaching* 18.6 (2023): 2051-2060.
- [23] Forcher, Leander, et al. "Is a compact organization important for defensive success in elite soccer?—Analysis based on player tracking data." *International Journal of Sports Science & Coaching* 19.2 (2024): 757-768.
- [24] Llana, Sergio, et al. "The right place at the right time: Advanced off-ball metrics for exploiting an opponent's spatial weaknesses in soccer." *Proceedings of the 14th MIT Sloan Sports Analytics Conference*. 2020.
- [25] Bojinov, Iavor, and Luke Bornn. "The pressing game: Optimal defensive disruption in soccer." *10th MIT Sloan Sports Analytics Conference*. Vol. 11. No. 12.03. 2016.
- [26] Javier Fernandez-Navarro et al. "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." *Frontiers in Psychology* 10 (2020): 3040.
- [27] Benedict Low et al. "The porous high-press? An experimental approach investigating tactical behaviours from two pressing strategies in football." *Journal of Sports Sciences* 39.19 (2021): 2199-2210.
- [28] Leander Forcher et al. "The keys of pressing to gain the ball - Characteristics of defensive pressure in elite soccer using tracking data", *Science and Medicine in Football* 8(2), 161-169, 2022
- [29] Forcher, Leander, et al. "The Success Factors of Rest Defense in Soccer—A Mixed-Methods Approach of Expert Interviews, Tracking Data, and Machine Learning." *Journal of Sports Science & Medicine* 22.4 (2023): 707.
- [30] Mat Herold et al. "Off-ball behavior in association football: A data-driven model to measure changes in individual defensive pressure." *Journal of Sports Sciences* 40.12 (2022): 1412-1425.
- [31] Leander Forcher et al. "Prediction of defensive success in elite soccer using machine learning-Tactical analysis of defensive play using tracking data and explainable AI." *Science and Medicine in Football* (2023): 1-16.
- [32] Chaoyi Gu et al. "Player Pressure Map - A Novel Representation of Pressure in Soccer for Evaluating Player Performance in Different Game Contexts" *Proceedings of the 18th MIT Sloan Sports Analytics Conference*. 2024
- [33] Pascal Bauer and Gabriel Anzer. "Data-driven detection of counterpressing in professional football: a supervised machine learning task based on synchronized positional and event data with expert-based feature extraction." *Data Mining and Knowledge Discovery* 35.5 (2021): 2009-2049.
- [34] Hudl Statsbomb, <https://statsbomb.com/>
- [35] Javier Fernández and Luke Bornn. "Soccermap: A deep learning architecture for visually-interpretable analysis in soccer." *Machine Learning and Knowledge Discovery in Databases. Applied Data Science and Demo Track: European Conference, ECML PKDD 2020, Ghent, Belgium, September 14–18, 2020, Proceedings, Part V*. Springer International Publishing, 2021.

Appendix A

A.1. Description of features used for model training

Table 3 describes the detailed overview of features used for model training. The Hudl Statsbomb data catalog provides detailed definitions of each action and its success criteria.

Category	Description
Action Type	The type of each action, with a total of 23 types including pass, cross, throw-in, crossed-freekick, short-freekick, crossed-corner, short-corner, take-on, foul, tackle, interception, shot, penalty-shot, freekick-shot, keeper save, keeper punch, keeper claim, keeper pick-up, clearance, bad touch, dribble, goal kick, pressing.
Result	The outcome of each action, with six types including fail, success, offside, own goal, yellow card, and red card.
Body Part	The body part used to perform each action, with six types including foot, left foot, right foot, head, head/other, and other.
Time	The time when each action was performed, measured in seconds.
Location	The coordinates where each action started and ended. The coordinate system is defined on a full-pitch scale of 105 x 68 meters, following FIFA's recommended field of play dimension.
Polar	The polar coordinates where each action started and ended, calculated as the angle relative to the center of the opponent's goal.
Movement	The distance covered by each action, calculated as the distance between the start and end locations.
Possession Status	Indicates whether the team performing the current action is the same as the team from the previous action. (Boolean variable)
Time Delta	The time difference (in seconds) between the current and previous action.
Space Delta	The distance covered between the locations of the current and previous actions.
Goal Score	The number of goals scored immediately by each team after the action.
Relative Distance	The distances from the start and end locations to the sideline and goal line.
Angle	The angle between the start and end locations of the action.
Speed	The speed of the ball during the action, calculated as the distance between the start location of the current action and the end location of the previous action, divided by the time difference between them.

Distance to Opponent	The distance to the nearest opponent player. If 360 data is not available, the feature is not used.
Closest Players	The locations of the n closest teammates and opponents relative to the action's location. If 360 data is not available, the feature is not used.

Table 3. Category and description of features used for model training. This table lists and describes the comprehensive set of features used for training the model, as defined in the Hudl StatsBomb data catalog. Each feature is precisely defined to capture both the technical execution and contextual dynamics of football actions, ensuring robust and accurate modeling of game scenarios.

A.2. Implementation of SoccerMap

SoccerMap [35] is a CNN-based framework that predicts probability surfaces of potential passes by generating low-level inputs. The input data is transformed into a 68×105 grid, where each cell represents spatial information about players. Player positions are encoded into separate channels for the presser, teammates, and opponents, along with additional features capturing the distance and angle between each cell and specific reference points (e.g., the presser and the opponent's goal). The final input, structured as number of channels $\times 68 \times 105$, is processed through convolutional layers to extract spatial features and predict the probability of successful pressing. The CNN architecture consists of five convolutional layers and four max-pooling layers, with replication padding applied to each convolutional layer. The output of the final convolutional layer is flattened and passed through a linear function. Additionally, historical event data is incorporated by including probability maps with the same feature structure as input, ensuring contextual information is captured. Unlike conventional machine learning models, SoccerMap imposes no constraints on the order of input features or the number of players for interpolation, making it highly adaptable to diverse soccer scenarios. The model is implemented in PyTorch and trained using Cross Entropy loss with a batch size of 16, a learning rate of $1e-3$, and a weight decay of $1e-8$. The step learning rate scheduler is used to adjust the learning rate, and early stopping is applied with a patience value of 10.

A.3. Code Availability

All code and associated materials used in this study are publicly available in a GitHub repository at <https://github.com/leemingo/sr-press>. The repository includes scripts for data preprocessing, model training, evaluation, and application, along with configuration files and a comprehensive README that provides detailed instructions on setting up the environment and executing the experiments.