GPS - A Metric for Evaluating Goalkeeper Positioning

Lou Zhou[1]
[1]Department of Sport Management, Rice University, Houston, TX

Abstract

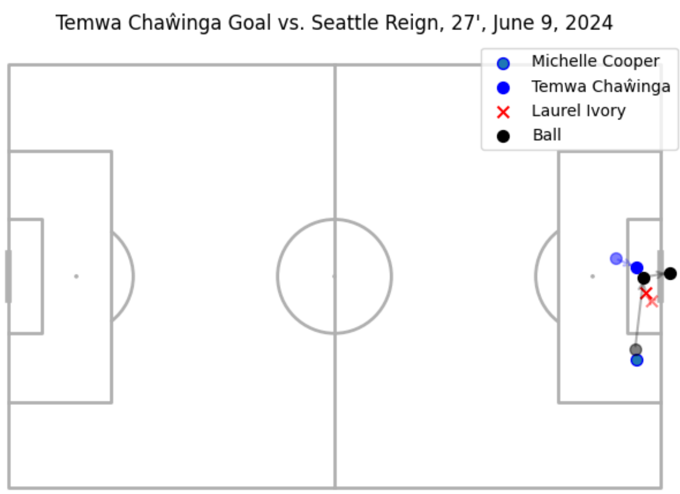
The positioning of the goalkeeper is central to defending threatening situations. Effective positioning can turn high-pressure moments into manageable saves and allow the goalkeeper to disrupt dangerous passes. Conversely, poor positioning can convert routine saves into goals and open pathways that make conceding inevitable. Despite this importance, most existing goalkeeper metrics focus primarily on a goalkeeper's shot-stopping ability. Therefore, we propose Goalkeeper Position Score (GPS), a metric used to evaluate goalkeeping positioning over the course of a threatening play. With this metric, we both calculate the defensive value that the goalkeeper generates from taking a certain position and contextualize that value compared to other expected locations. Using 2024 NWSL and 2024-2025 WSL tracking data, we then applied this metric to a real match scenario and evaluated goalkeeper performance across thousands of threatening game states. Our results show that GPS provides meaningful insight into quantifying the positioning abilities of goalkeepers, offering a complementary perspective to goalkeeping effectiveness outside of shot-stopping.

1 Introduction

Defensively, the positioning of the goalkeeper is crucial in preventing goals when the opposing team is attacking. Being able to anticipate and position oneself properly prior to a shot or developing play can turn a threatening play into a simple save. Conversely, poor positioning can turn routine saves into a conceded goal. No matter the goalkeeper's shot-stopping ability, the positioning of the player is key to prevent goals.

With the importance of this positioning, current metrics like save percentage and differences between expected goals and actual goals allowed evaluate goalkeeper performance primarily on shot-stopping ability and do not adequately assess the quality of goalkeeper positioning during threatening situations. Looking at these metrics, we only get a view of a goalkeeper's effectiveness in stopping shots, but we gain little insight into whether the goalkeeper was well-positioned to make the save in the first place.

For example, consider the goal where Temwa Chaŵinga scored against Seattle Reign during the 2024 NWSL season. In this situation, KC Current forward Michelle Cooper is able to break free and find space on the right wing, playing a pass across the face of the goal for an easy finish for Chaŵinga.



However, given the cross's close proximity to the goal, this situation naturally raises a key question: was the goalkeeper, Laurel Ivory, optimally positioned to prevent the goal? If Ivory had adopted a more advantageous position, could she have been able to better position herself to either prevent this pass or make Chaŵinga's finish more difficult while preventing a direct shot from Cooper? Alternatively, was her positioning already near optimal, making the goal effectively unavoidable given the 2 vs. 1 situation?

Given these questions, this paper introduces a novel metric, Goalkeeper Positioning Score (GPS), to evaluate and quantify the quality of goalkeeper positioning in soccer during threatening situations. While there has been some work regarding goalkeeper positioning from both Lamas et al. [5] and Gottini [4], these works involve only shot situations where works look to find the expected or optimal location for the goalkeeper for a shot. Here, we propose a new metric which can be applied to all threatening situations where the goalkeeper could have an impact on the play during that game state (e.g. collecting crosses or preventing easy passes leading to goals).

Using this framework, we can compute GPS, which describes the difference between the danger of the state with the actual goalkeeper's position and the expected danger across all locations. To compute these values, we derive two estimates: the probability of the goalkeeper being positioned at a certain location and the defensive value that the goalkeeper provides from being at a specific location.

As an implementation, we also build a machine-learning approach off previous positioning work to estimate these values on the 2023-2024 Women's Super League (WSL) and National Women's Soccer League (NWSL) seasons using full tracking data provided by SkillCorner and event data provided by Wyscout. Using our estimates, we can then apply our metric to determine Ivory's positioning quality during the goal as well as find goalkeepers who consistently positioned themselves optimally across NWSL during the 2024-2025 season.

The remainder of this work is organized accordingly: Section 2 presents our metric for quantifying the strength of a goalkeeper's position given the game state, Section 3 presents the models built to quantify the estimates used to compute GPS, Section 4 presents insights in the application of GPS onto NWSL data, and Section 5 discusses limitations and potential directions for future work.

2 Computing Goalkeeper Positioning Score

In deriving GPS, we seek to build an interpretable metric which compares the danger, D, of the game state, S, given the observed positioning of the goalkeeper, to the expected danger across all potential goalkeeper locations. With this approach, we can isolate the marginal effect of the goalkeeper's positioning, revealing how the goalkeeper's position changes the danger of the state relative to a league-wide expectation.

By comparing to an expectation, we are also able to adjust for the underlying danger that the goalkeeper faces during a specific threatening game state. We would not punish a goalkeeper for standing in an inherently dangerous position if the situation were threatening independent of goalkeeper positioning. Likewise, we would not reward a goalkeeper for occupying a safer location if the situation itself were not very threatening, regardless of the positioning of the goalkeeper.

Furthermore, this approach allows us to normalize for situations where the optimal position is obvious and reward goalkeepers for finding better positioning than where we might expect other goalkeepers to position themselves.

We can then interpret GPS as the reduction (or creation) of danger during a threatening situation as a result of the goalkeeper's positioning compared to where we might expect the goalkeeper to be.

More formally, for a game state S, we define the GPS with goalkeeper position at location $GK = l$ as:

$$GPS(S|GK = l) = E_{GK|S}[D(S|GK)] - D(S|GK = l) \quad (1)$$

Here, a higher GPS indicates stronger positioning while a lower GPS indicates weaker positioning.

2.1 Computing Expected Danger

In deriving the expected danger across all locations, we build models to compute two quantities for a given goalkeeper location l'.

1. $D(S|GK = l')$, The inherent danger of the state S given goalkeeper being in location l'
2. $P(GK = l'|S)$, The league-wide probability of a goalkeeper being in location l' given game state S

Using these two quantities for all locations l', we can then compute the expected danger of the goalkeeper position given state S as, for all locations on the field $\mathcal{L}$:

$$E_{GK|S}[D(S|GK)] = \int_{\mathcal{L}} D(S|GK = l')P(GK = l'|S)dl'$$

In practice, since we use a discrete 105×68 array as our pitch surface, this value becomes

$$E_{GK|S}[D(S|GK)] = \sum_{l' \in \mathcal{L}} D(S|GK = l')P(GK = l'|S) \tag{2}$$

To calculate GPS, we also find the value of the goalkeeper standing at the observed position using the same model used to compute $D(S|GK = l')$.

3 Implementing GPS

3.1 Data

In implementing GPS, we use tracking data from SkillCorner, which describes the locations of all players and the ball on the field throughout the game, across both the 2024 NWSL and 2024-2025 WSL season. From these leagues, 265 matches were deemed usable by SkillCorner for further downstream analysis. Furthermore, to provide context behind the tracking data, we also use synced Wyscout event data which describes specific actions (e.g. shot or pass) throughout the course of a game.

Since we are only interested in threatening situations, we filter all frames so that we only include situations where the ball is in a location where the Expected Threat (xT) [7] is at least 0.05. With these criteria, we create a training dataset of 232,450 game states (frames) across both leagues where the possessing team has the ball in a dangerous situation. Furthermore, we extract a dataset of 151,408 game states from the NWSL season to evaluate the metric and generate results.

3.2 Estimating the Probability of a Goalkeeper's Position

To generate the probability surface for a goalkeeper, using the implementation from Robberechts et al. [6], we adapt the SoccerMap architecture from Fernandez and Bornn [2], which looks to predict the end location for potential passes. Here, instead of predicting end location for these passes, defined by a sparse matrix where a value of 1 is given for the observed pass destination, we look to predict the location of the goalkeeper, which is similarly defined by a sparse matrix with a value of 1 for the goalkeeper location.

To represent the game state, we build a tensor representation using tracking data with the following channels, adapted heavily from Robberechts et al. [6]:

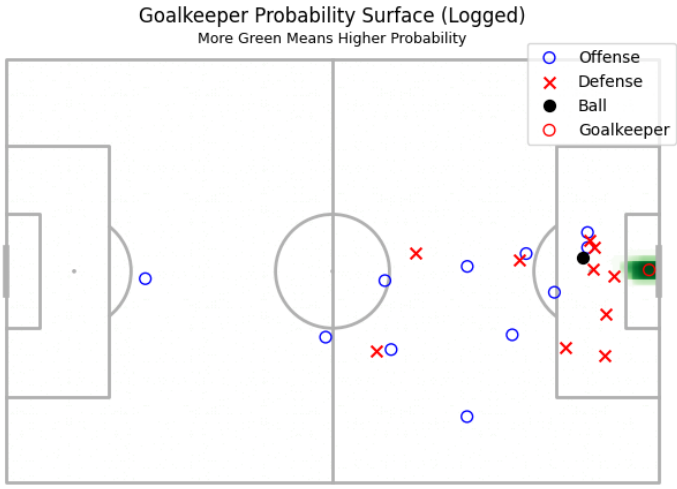
1. Two channels describing the locations of both the attacking and defending players
2. Two channels describing each location's distance from both the goal and ball
3. Two channels describing the sine and cosine between every location and the ball
4. One channel describing the angle between every location and the goal
5. Two channels describing the x and y velocity of the ball

With this architecture, we also add the following channels describing the velocity of players:

1. Two channels describing the x and y velocities of all attacking players
2. Two channels describing the x and y velocities of all defending players

Since we can get unstable estimates of velocity if we use the distances between the t and $t-1$ frame, we compute velocities at frame t by taking the displacement between the frames $t-5$ and t, then dividing by 0.5 seconds since the tracking data is captured at 10 frames per second.

We then tune the model's hyperparameters by using a grid search method and train the model with early stopping using a patience of 10 epochs. Using this model, we can then generate probability surfaces describing the likelihood of the goalkeeper being at a specific location across the entire field.



3.3 Estimating the Defensive Value of a Goalkeeper's Position

In estimating defensive value from positioning, we quantify the danger of the situation by looking at how likely the possession team will score in the next 10 seconds. Following previous work [6], we define this danger as the probability that at least one of the shots taken in that window results in a goal. More formally, let $\mathcal{P}(S)$ be the set of shots taken by the team in possession during state S in the next 10 seconds, and let xG_j be the expected goals value of shot j. We define $D(S)$ as

$$D(S) = 1 - \prod_{j \in \mathcal{P}(S)} (1 - xG_j)$$

By using this 10 second window, we can get a strong idea of the imminent danger associated with state S, capturing not only potential immediate shot quality but also the downstream threat that arises from the goalkeeper's positioning and the developing play. With this 10 second window, we also avoid situations in which the goalkeeper's position at S has little influence on events occurring much later.

Using the tracking data, we then use engineered features to build a meaningful feature set[1] which represents the game state. Besides obvious geometric features like ball velocities and distance from the goal, we include features such as:

1. Number of attackers in the box
2. Number of defenders within 3 meters from the goalkeeper
3. Defensive line depth

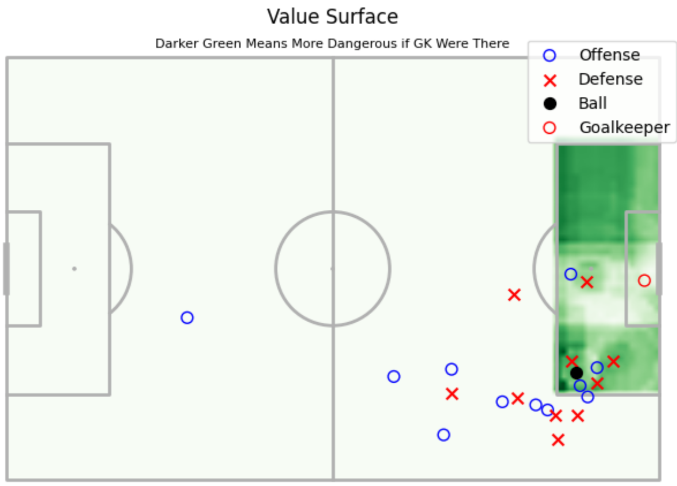
From there, we can then predict $D(S|l')$ with an XGBoost [1] approach to derive a value for the goalkeeper's positioning at l', where a higher predicted value means a more dangerous situation.

Using this model, following a random search hyperparameter tune, we gain an out-of-sample RMSE of 0.046 on our strongest parameter set. While performing stronger than the naive baseline of simply predicting the global mean $D(S)$, giving us an RMSE of 0.092, it would still be useful to apply other approaches to see if there are stronger models for this prediction task.

To find the expected danger across all locations, because of computational complexity, we compute these predictions across hypothetical goalkeeper positions only across the entire 18-yard box, as, unlike the goalkeeper probability model, for each hypothetical position, we would need to recompute many more features relating to the position of the goalkeeper, increasing computational complexity.

[1] The full feature set can be found in this work's GitHub Repository

Since we exclusively look at threatening situations, as except for very rare edge cases, we do not expect the goalkeeper to be outside the 18-yard box, we only compute within this 18-yard box.



Using this surface across the entire box, we can then combine with our probability surface to derive the expected danger of the game state using (2). We then find the difference between this expected danger and the actual danger from the goalkeeper's position, generating GPS.

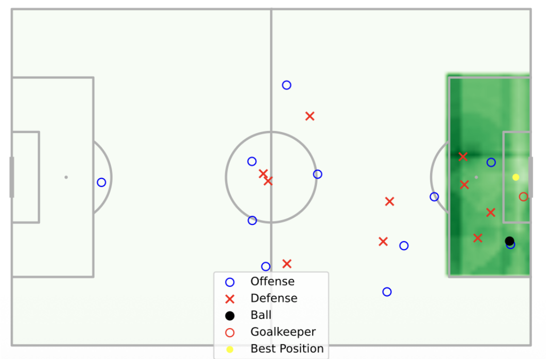
4 Applications on NWSL Data

4.1 Applying GPS to Chaŵinga's Goal

In applying GPS to Chaŵinga's goal, we can both gauge the strength of OL Reign Goalkeeper Laurel Ivory and contextualize this strength in comparison to where we might expect Ivory to position herself.

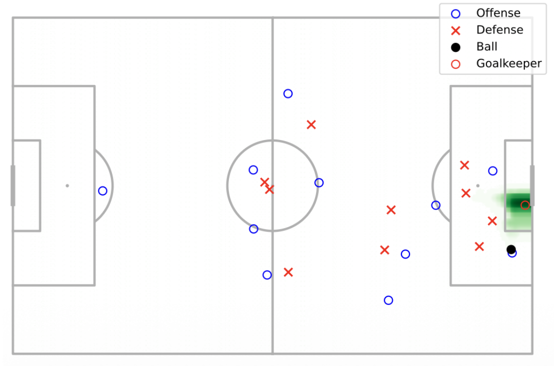
Throughout the course of the play, Ivory has an average GPS of -0.075, indicating that, on average throughout all frames of the play, Ivory's positioning increased the probability of a goal during the state by 7.5% compared to expected goalkeeping positions. This indicates that, over the duration of the play, she positioned herself less effectively than would be expected given the evolving game state.

Looking specifically at the frame where Cooper makes the pass to Chaŵinga, the danger surface for this moment identifies the optimal defensive position as being further off the near post, anticipating the incoming pass rather than the low-probability direct shot from Cooper's angle.



This recommendation aligns with the underlying tactical scenario. Although standing further off the near post can increase the chance of conceding a direct shot from Cooper, that shot is low-quality relative to the near-certain finish available to Chaŵinga if the pass arrives uncontested. The optimal position therefore prioritizes defending the cross.

Through the GPS metric, we can also contextualize this positioning by comparing the observed position with a probability surface of other positions.



Based on this surface, while Ivory positions herself in an area commonly occupied by other goalkeepers, the probability surface highlights the fact that we could also expect that the goalkeeper would move slightly farther off the near post in this situation. This region overlaps closer with the model's optimal positioning, suggesting that Ivory's choice was typical but not maximally effective given the particular threat posed by Chaŵinga's run, hence the weak GPS value.

Taken together, this moment highlights how GPS allows us to interpret Ivory's positioning in the context of both optimal and expected behavior. While her location is not an obvious positional error, the model reveals that a small adjustment that other goalkeepers might make can materially alter the scoring landscape when defending a dangerous cross. In this case, by shifting just a few steps further off the post, Ivory would have been able to reduce the likelihood of conceding the final shot by better protecting the central passing lane without substantially increasing the threat of a near-post attempt by Cooper.

More broadly, this example demonstrates how GPS can diagnose the small positioning decisions goalkeepers make in threatening sequences, quantifying how each positional adjustment affects the attacking team's expected goal probability. The GPS metric allows for the quantification of goalkeeper positioning, contextualized by where we might expect other goalkeepers to position themselves.

4.2 Best Positioned Goalkeepers

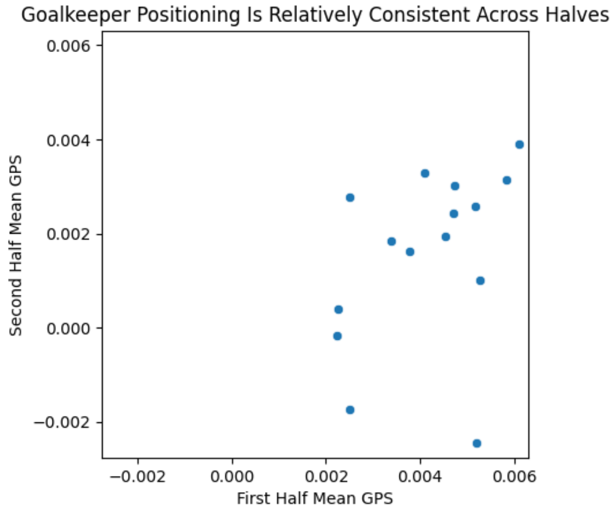
By finding the average GPS for each goalkeeper in the NWSL, we can get a strong intuition about which goalkeepers consistently position themselves in better locations than other goalkeepers. As shown in Table 1, several goalkeepers separate themselves from the league in terms of average positional value. Katie Lund, a finalist for the 2023 NWSL Goalkeeper of the Year, and AD Franch, a two time winner of the award, appear at the top of the rankings, reflecting their ability to consistently maintain advantageous positioning compared to their peers. Their elevated GPS values highlight a pattern of consistently occupying positions that reduce expected scoring compared to the league-wide baseline.

Table 1: Top Goalkeepers by Mean GPS, Minimum 1000 Frames (NWSL 2024)

Name	Mean GPS	Frames
K. Lund	0.0057	10504
A. Franch	0.0053	6321
A. Anderson	0.0052	1141
C. Miller	0.0050	1847
L. Proulx	0.0049	2468

The remaining goalkeepers in the top five, Anderson, Miller, and Proulx, also register strong average GPS values, though their lower frame counts introduce potential variability in interpreting their performance. While their positioning is strong on a per-frame basis, the smaller sample sizes mean that these estimates are likely more volatile and potentially influenced by potential variations in game states that they might face. In addition, with mean GPS values for all NWSL goalkeepers falling close to 0 and GPS values per frame tending to fall within (-0.01, 0.02), there are no goalkeepers who are consistently deviating from the expected positions.

Looking at the stability of GPS throughout the season, we also note that the correlation between 1st half and 2nd half GPS is relatively strong, with $r = 0.40$.



This stability indicates that goalkeeper positioning behaviors are not purely situational or random but instead reflect consistent underlying tendencies that persist across matches. This consistency suggests that GPS captures meaningful patterns in how goalkeepers read and respond to developing attacks, rather than frame-by-frame variance driven by noise or isolated plays. However, with a small sample size ($n = 15$), it will be useful to calculate GPS in future seasons because larger samples will be needed to more confidently assess the robustness of this relationship.

Nevertheless, it should be noted that GPS differentiates goalkeepers not by their shot-stopping ability, but by the quality of their preparatory positioning leading up to shots and dangerous passes. A player might have a strong GPS value as they tend to position themselves optimally, but their weak shot-stopping ability might make them a weaker overall goalkeeper. Conversely, some goalkeepers may post lower GPS values despite strong overall performance, as elite shot-stopping ability can compensate for suboptimal positioning. In this way, GPS and shot-stopping metrics capture different dimensions of goalkeeping effectiveness.

5 Discussion

With the importance of positioning in defending threatening situations, this work introduces GPS, a metric for quantifying and evaluating goalkeeper positioning. To achieve this, we decompose goalkeeper positioning into (1) where we would expect the goalkeeper to be given the game state and (2) how dangerous the game state would be when a goalkeeper occupies a specific location. With this decomposition, we are able to build an interpretable measure which both captures the strength of the goalkeeper's position and contextualizes this positioning strength in comparison to expected goalkeeper positioning.

Using the example of Chaŵinga's goal, GPS is able to diagnose how positioning choices might significantly alter the threat of a specific game state. By considering both the value of the goalkeeper's positioning as well as the expected location of other goalkeepers, GPS is able to normalize for the actual threat of the game state and isolate the portion of that danger attributable specifically to the goalkeeper's positional decision. This allows us to understand not only whether the goalkeeper's positioning increased or reduced the likelihood of conceding, but also whether their choice was meaningfully different from what we would expect an average goalkeeper to do.

More broadly, GPS allows us to evaluate goalkeepers outside of simple shot-stopping ability. Because a large part of goalkeeping performance involves these actions outside of shots, like collecting crosses, GPS was designed to assess the impact of these early decisions on a developing attack. By measuring how a goalkeeper's location influences the evolving scoring landscape during a threatening game state, GPS captures the proactive elements of goalkeeping that traditional shot-based metrics cannot detect, providing a more holistic understanding of a goalkeeper's defensive value.

5.1 Future Work and Limitations

Regarding our implemented models which calculated GPS, several limitations highlight opportunities for further refinement. First, both our defensive value model and goalkeeper position probability model operate only on information from a single frame. This limitation can be problematic as it ignores the context of how the current game state arose. For instance, a sudden turnover leading to an immediate counter-attack may leave the goalkeeper in a naturally compromised or unreachable position. Without accounting for this buildup, the current models may unintentionally penalize goalkeepers for occupying locations that are suboptimal yet unavoidable given the sequence of preceding events.

Looking at the value model, there is always an opportunity to expand the feature set of our gradient boosting model so that a richer representation of the game state could be created. For example, features like player-specific profiles might be useful as positioning may be affected by if the attacking players are left or right footed. These features could help the current model capture a stronger estimate of the true defensive value associated with different locations. Likewise, with more computational power, it would be useful to predict over the entire field as opposed to just the box to account for rare scenarios where being outside the box might be optimal.

Furthermore, in using Expected Threat to filter for threatening situations, we can unintentionally restrict our analysis to danger defined solely by the location of the ball. This definition can inadvertently remove dangerous situations where the ball might not be close to the goal. For example, if the offense is through on goal but not necessarily close to the goal, the goalkeeper's decision to either rush out and meet the attacking player or stay in the goal is important in defending this situation. Using the current xT metric, we would not consider these plays due to the position of the ball. In the future, it may be useful to use a metric which also considers the locations of players, like expected possession value [3].

Acknowledgements

The author would like to thank Scott Powers for his valuable feedback and guidance throughout the development of this work. Appreciation is also extended to the authors of un-xPass: Measuring Soccer Player's Creativity [6] for making their SoccerMap implementation publicly available, which served as an essential foundation for parts of this project. Finally, the author gratefully acknowledges the American Soccer Insights Summit and SkillCorner for providing the tracking and event data that made this work possible.

References

[1] 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*, KDD '16, page 785–794, New York, NY, USA, 2016. Association for Computing Machinery.

[2] J. Fernandez 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: European Conference, ECML PKDD 2020, Ghent, Belgium, September 14–18, 2020, Proceedings, Part V*, page 491–506, Berlin, Heidelberg, 2020. Springer-Verlag.

[3] J. Fernandez and L. Bornn. A framework for the fine-grained evaluation of the instantaneous expected value of soccer possessions. *Mach. Learn.*, 110(6):1389–1427, June 2021.

[4] G.-A. Gottini. Quantitative analysis of football goalkeeper positioning. Master's thesis, ETH Zurich, 2023.

[5] L. Lamas, R. Drezner, G. Otranto, and J. Barrera. Analytic method for evaluating players' decisions in team sports: Applications to the soccer goalkeeper. *PLoS One*, 13(1), 2018.

[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*, KDD '23, page 4768–4777, New York, NY, USA, 2023. Association for Computing Machinery.

[7] K. Singh. Introducing expected threat (xt), 2019.