

VAEP: An Objective Approach to Valuing On-the-Ball Actions in Soccer (Extended Abstract)

Tom Decroos¹, Lotte Bransen², Jan Van Haaren^{1,2} and Jesse Davis¹

¹KU Leuven, Department of Computer Science & Leuven.AI, Belgium

²SciSports, Netherlands

{tom.decroos, jesse.davis}@kuleuven.be, {l.bransen, j.vanhaaren}@scisports.com

Abstract

Despite the fact that objectively assessing the impact of the individual actions performed by soccer players during games is a crucial task, most traditional metrics have substantial shortcomings. First, many metrics only consider rare actions like shots and goals which account for less than 2% of all on-the-ball actions. Second, they fail to account for the context in which the actions occurred. This work summarizes several important contributions. First, we describe a language for representing individual player actions on the pitch. This language unifies several existing formats which greatly simplifies automated analysis and this language is becoming widely used in the soccer analytics community. Second, we describe our framework for valuing any type of player action based on its impact on the game outcome while accounting for the context in which the action happened. This framework enables giving a broad overview of a player's performance, including quantifying a player's total offensive and defensive contributions to their team. Third, we provide illustrative use cases that highlight the working and benefits of our framework.

1 Introduction

A fundamental question for soccer clubs is:

How will a soccer player's actions impact his or her team's performances in games?

Understanding the answer to this question informs crucial tasks such as player acquisition, player evaluation, and scouting. Typically, this question was addressed in a subjective manner via video analysis by human scouts. The availability of new data sources such as optical tracking and event stream data has spurred interest in automated approaches for objectively quantifying the impact of the individual actions performed by a player during a match.

⁰The original paper of this extended abstract was presented at the 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2019) [Decroos *et al.*, 2019].

Unfortunately, the extremely low-scoring and dynamic nature of soccer matches makes this a difficult task. Most existing approaches for valuing actions in soccer suffer from three important limitations. First, the vast majority of these approaches focus on the concept of the expected value of a goal attempt [Lucey *et al.*, 2014; Caley, 2015; Altman, 2015; Mackay, 2017], which, while more frequent than goals, still only occur rarely. Second, existing approaches ignore much of the action's context, that is, the circumstances under which the action was performed. For example, some approaches only consider the action's location [Singh, 2019; Van Roy *et al.*, 2020] and ignore important factors such as the goal difference, time remaining, and speed of play (e.g., if a pass occurs during a rapid counter attack). Third, most approaches only consider immediate effects and fail to account for an action's effects a bit further down the line.

This extended abstract of our KDD'19 paper [Decroos *et al.*, 2019] describes our data-driven framework for valuing actions in a soccer game. Unlike most existing work, it considers all types of actions (e.g., passes, crosses, dribbles, take-ons, and shots) and accounts for an action's context as well as its possible longer-term effects. Intuitively, an action value reflects the action's expected influence on a match's scoreline. That is, an action valued at +0.05 is expected to contribute 0.05 goals in favor of the team performing the action, whereas an action valued at -0.05 is expected to yield 0.05 goals for their opponent. This extended abstract summarizes five important contributions:

1. A language for representing player actions;
2. A framework for valuing player actions;
3. A model for predicting short-term scoring and conceding probabilities at any moment in a game;
4. A number of use cases showcasing our most interesting results and insights; and
5. A Python package¹ that implements the first four contributions and is becoming widely used in the soccer analytics community.

2 SPADL: A Language for Describing Actions

Event stream data annotates the times and locations of specific events (e.g., passes, shots, and cards) that occur in

¹<https://github.com/ML-KULeuven/socceraction>

a game. Vendors such as Opta, Wyscout, and StatsBomb collect this type of data by having humans annotate match videos. Three key challenges arise when analyzing this data.

Challenge 1: Event stream data serves multiple different objectives (e.g., reporting information to broadcasters, analysts at soccer clubs). Hence, the data is not necessarily designed to facilitate analysis. Some important information is missing (e.g., Wyscout does not record exact end locations for shots) or irrelevant information is included (e.g., Opta records weather changes).

Challenge 2: Each vendor of event stream data uses their own unique terminology and definitions to describe the events that occur during a game. Hence, software written for event stream data has to be tailored to a specific vendor and cannot be used without modifications to analyze data from another vendor.

Challenge 3: Most vendors offer optional information snippets per type of event. For example, Opta often specifies details about the exact type of foul that was committed. While sometimes useful, this information makes it extremely hard to apply automatic analysis tools.

The need for a uniform representation language is a typical AI problem. To this end, we propose SPADL (Soccer Player Action Description Language), a language to describe soccer actions that unifies event stream formats from different vendors (e.g., WyScout, StatsBomb, Opta) into a common vocabulary that enables subsequent data analysis. SPADL describes player *actions*, as opposed to the formats by commercial vendors that describe *events*. The distinction is that actions are a subset of events that require a player to perform the action. For example, a passing event is an action, whereas an event signifying the end of the game is not an action.

Unlike other event stream formats, SPADL always stores the same eight attributes for each action: time, start location, end location, involved player, team of the involved player, the action type (e.g., *pass*, *shot*, *dribble*), the body part used to perform the action (e.g., *foot*, *head*), and the result of the action (e.g., *success* or *fail*). Storing the same eight attributes for each action addresses the third challenge, as excluding optional information snippets enables us to more easily apply automatic analysis tools.

3 VAEP: A Framework for Valuing Actions

A fundamental task in soccer analytics is attempting to understand the value of each action that a player performs during a game. This task can be formally defined as follows:

Given: An on-the-ball action a_i ;

Do: Learn a function that assigns a value $V(a_i)$ to the action.

The effect of every on-the-ball action is that it alters the game state. That is, an action a_i moves the game from game state S_{i-1} to game state S_i . VAEP leverages this by measuring the value of this change in game state:

$$V(a_i) = V(S_i) - V(S_{i-1})$$

where $V(\cdot)$ represents the value of a game state or action. Hence, the primary question becomes: How should we value a game state?

Intuitively, a helpful action is one that benefits a player's team, either by increasing the chance that his team scores or decreasing the chance that the opposing team scores. These actions should be positively valued. In contrast, actions that do the opposite hurt the player's team and should be negatively valued. This observation suggests that one way to value a game state is:

$$V(S_i) = P_{scores}^k(S_i) - P_{concedes}^k(S_i)$$

where $P_{scores}^k(S_i)$ ($P_{concedes}^k(S_i)$) is the probability that the team possessing the ball in state S_i will score (concede) within the next k actions. Typically, k is a small number like five or ten as the effects of most actions are temporally limited. Practically, the task then becomes estimating these two probabilities for each state, which can be solved by training a machine learning model that predicts a probability.

The VAEP framework provides a simple approach to valuing actions that is independent of the representation used to describe the actions. The framework's strength is that it transforms the subjective task of valuing an action into the objective machine learning task of predicting the likelihood of a future event in a natural way.

4 Estimating Scoring and Conceding Probabilities

This section describes our method for estimating the scoring and conceding probabilities required by the VAEP framework. This task can be formally described as follows:

Given: A game state S_i ;

Do: Estimate $P_{scores}^k(S_i)$ and $P_{concedes}^k(S_i)$.

For both binary classification problems we train a probabilistic classifier to estimate the probabilities. In principle, any machine learning model (e.g., logistic regression, random forest, neural network) that predicts a probability could be used to address these tasks. Empirically, we found that gradient boosted decision tree ensembles such as XGBoost [Chen and Guestrin, 2016] and CatBoost [Prokhorenkova *et al.*, 2018] work well in practice.

Applying a standard machine learning algorithm requires converting the sequence of actions $[a_1, a_2, \dots, a_m]$ describing an entire game into examples in a feature-vector format. Thus, one training example is constructed for each game state S_i . We compute 151 features that impact the probability of a goal being scored in the near future. These features are based on the last three actions of a game state and can be divided into three categories.

First, the *SPADL features* are categorical and real-valued features based on information explicitly included in the SPADL representation. We consider categorical features for an action's type, result, and the body part used by the player performing the action. Similarly, we consider real-valued features for the (x, y) -coordinates of the action's start and end locations, and the time elapsed since the start of the game.

Second, the *complex features* combine information within and across consecutive actions. Within each action, these features include (1) the distance and angle to the goal for both

	TIME	PLAYER	ACTION	P_{scores}	VALUE
○	1 92m4s	S. Busquets	pass	0.03	0.00
○	2 92m6s	L. Messi	pass	0.02	- 0.01
●	3 92m8s	S. Busquets	pass	0.03	+ 0.01
- -	4 92m11s	L. Messi	take on	0.08	+ 0.05
●	5 92m12s	L. Messi	pass	0.17	+ 0.09
■	6 92m14s	A. Vidal	shot	1.00	+ 0.83

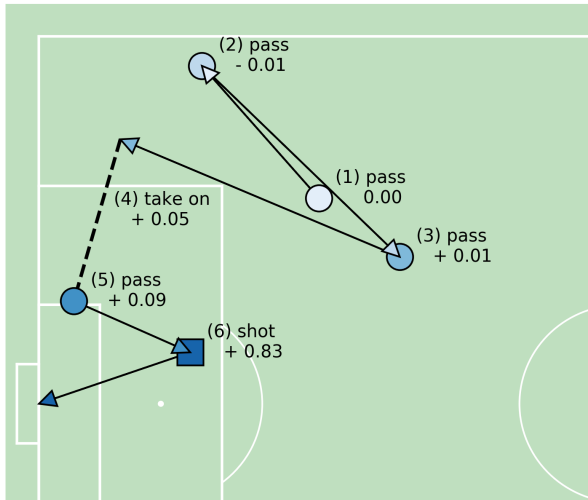


Figure 1: The attack leading up to Barcelona’s final goal in their 3-0 win against Real Madrid on December 23, 2017.

the action’s start and end locations, and (2) the distance covered during the action in both the x and y directions. Between two consecutive actions, we compute the distance and elapsed time between them and whether the ball changed possession. These features provide some intuition about the current speed of play.

Third, the *game context features* include the scoreline and goal difference after action a_i . This information is relevant because scoring probabilities depend on these factors [Robberechts *et al.*, 2019].

5 Experiments

Evaluating our framework is challenging as no objective ground truth action values or player ratings exist. Lack of ground truth data is an important challenge that frequently occurs when applying AI techniques to real-world data, but is not often discussed in traditional AI papers. In this extended abstract, we provide intuitions into how our framework behaves, show the natural trade-off between the quantity and quality of players’ actions, and list the top-10 players of the 2019/2020 English Premier League season. The full paper contains a number of other use cases [Decroos *et al.*, 2019].

5.1 Intuition behind the Action Values

Figure 1 illustrates how our framework works by visualizing the actions and their corresponding values that led to Barcelona’s goal in the 93rd minute of their away game

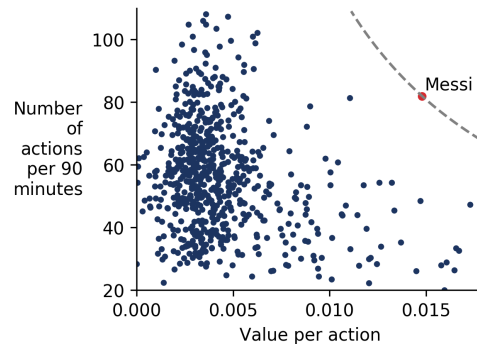


Figure 2: Scatter plot of players in the 2017/2018 season who played at least 900 minutes in the Spanish or English league. The plots contrast the average number of actions performed per 90 minutes with the average value of the actions of the player. As shown by the grey-dotted isoline, Lionel Messi is clearly in a class of his own.

against Real Madrid on December 23, 2017. The attack consists of six actions. The first three actions receive relatively neutral action values (i.e., 0.00, -0.01, +0.01) as they appear further away from the goal and thus do not have a strong effect on Barcelona’s scoring odds. The fourth action is more interesting; Messi receives the ball and dribbles past a Real Madrid defender into the box, which receives an action value of +0.05 for raising the scoring odds from 0.03 to 0.08.

The fifth action is a backwards pass that illustrates both Messi’s genius and the power of VAEP in recognizing this genius. Simple metrics often do not reward backwards passes as they move the ball further away from the goal. In this case however, Messi’s backwards pass is a great action as it moves the ball away from the crowded six yard box and sets up a shot with a better angle for a teammate. In a purely data-driven way, VAEP recognizes that this action increases the scoring odds from 0.08 to 0.17 and awards it a value of +0.09. Finally, Alex Vidal shoots and scores. For converting a 0.17 scoring chance to a goal, our framework rewards Vidal with an action value of +0.83. If Vidal had missed his shot, he would have been penalized with an action value of -0.17.

5.2 Trading Off Action Quality and Quantity

A natural tension exists between the quality and quantity of actions. If a player performs a high number of actions, then it is harder for each action to have a high value. Figure 2 shows the number of actions that players execute on average per 90 minutes (*quantity*) and the average value of these actions (*quality*) for those players who played at least 900 minutes during the 2017/2018 season in the Spanish and English leagues. The grey-dotted isoline shows the gap in VAEP rating between top-ranked Lionel Messi and the rest. The isoline is curved as a player’s rating is obtained by multiplying the average value per action (x -axis) and the average number of actions (y -axis). As shown by the isoline and more traditional statistics,² Messi is clearly in a class of his own.

²<https://fivethirtyeight.com/features/lionel-messi-is-impossible/>

Rank	Player	Team	Goals/90	Assists/90	VAEP/90
1	R. Mahrez	Manchester City FC	0.42	0.42	0.89
2	K. De Bruyne	Manchester City FC	0.31	0.59	0.79
3	S. Mané	Liverpool FC	0.57	0.24	0.69
4	S. Agüero	Manchester City FC	0.97	0.12	0.68
5	P. Aubameyang	Arsenal FC	0.62	0	0.62
6	D. Ings	Southampton FC	0.63	0.04	0.62
7	Mohamed Salah	Liverpool FC	0.61	0.08	0.61
8	J. Vardy	Leicester City FC	0.72	0.19	0.59
9	R. Sterling	Manchester City FC	0.46	0.04	0.59
10	Richarlison	Everton FC	0.36	0.07	0.56

Table 1: The top-10 players who played at least 900 minutes in the 2019/2020 English Premier League season up until April 2020 in terms of our VAEP player ratings.

5.3 New Results about the 2019/2020 Season

Table 1 shows the top-10 players in the 2019/2020 English Premier League season (up until April 2020) in terms of our VAEP player ratings. As expected, the list is dominated by attackers from Liverpool and Manchester City, the two best teams in the Premier League right now. A more surprising name in the list is Southampton striker Danny Ings. While he has suffered more than his fair share of injuries in his career, he has avoided the injury bug this season and has been in sensational form.³

6 Discussion and Perspectives

This paper falls within a growing line of data science research for analyzing sports data (e.g., [Liu and Schulte, 2018; Routley and Schulte, 2015; Decroos *et al.*, 2018; Pappalardo *et al.*, 2019; Bransen *et al.*, 2019; Fernández *et al.*, 2019; Rudd, 2011]) which has been fueled by the availability of data such as

- event stream data, which describes specific events that occur during a match;
- tracking data, which records the position of each player multiple times per second;
- athlete monitoring data, which records information such as heart rate and accelerations.

These data and the questions posed by domain experts (e.g., clubs, sports scientists, biomechanists) yield an incredibly rich and broad set of problems from an AI perspective. The following non-exhaustive list includes some illustrative examples. Pattern mining [Decroos *et al.*, 2018] and non negative matrix factorization [Decroos and Davis, 2019] have been used to understand tactics and playing style. Reinforcement learning and Markov decision processes have been used to analyze technical performance in a match [Liu and Schulte, 2018; Schulte *et al.*, 2017]. Multi-agent learning has been used to learn the value of teamwork from outcomes of pass interactions between players [Beal *et al.*, 2020]. Predictive modeling has been used to understand and manage training load [Jaspers *et al.*, 2018]. Deep learning [Kim *et al.*, 2018] and integer linear programming [Shitrit *et al.*, 2011] have been used for tracking players during a match.

³<https://www.theguardian.com/football/who-scored-blog/2020/jan/15/danny-ings-renaissance-england-southampton-premier-league>

Similarly, on a technical level these problems touch upon a number of core areas within AI:

Representation Sports data of various sources always needs to be preprocessed and put into a format compatible with AI techniques. For example, SPADL is a great way to represent soccer event stream data when training a classifier, but a more suitable representation is needed when applying a different AI technique such as pattern mining [Decroos *et al.*, 2018].

Interpretability Ultimately, experts are interested in translating the findings arising from analytics into practice. For example, an effect of the insights from expected goals models is that the number of long range shots have dramatically decreased in recent years.⁴

Decision making Decision making plays a crucial role in sports, and AI can often offer support. E.g., What strategy should a coach employ to maximize the odds of winning a game? Which action should a player perform in a specific situation? How long should an athlete train for?

Understanding behavior Professional athletes exhibit specific behaviors on the pitch that are often a combination of their innate knowledge of the sport, their role within their team, and their own personality. As sports data often essentially describes the actions of players, this is a great data source to understand their behavior. For example, [Bransen *et al.*, 2019] aim to understand how professional soccer players adapt their behavior to high mental pressure situations.

Experimental evaluation Often, real-world data lack ground truth labels. These may be hard to obtain or simply not exist. This is frequently the case in sports analytics as there are no objectively best players. Moreover, a team’s tactical plan is only known to the team itself. However, the richness of the sports domain encourages researchers to devise creative solutions for this problem. Some examples are measuring correlation with proxy signals such as player salary [Liu and Schulte, 2018] and evaluating by asking domain experts [Bransen *et al.*, 2019; Pappalardo *et al.*, 2019]).

For all of these reasons, we believe that sports will remain a fruitful source of real-world problems and a test bed for AI in the coming years.

Acknowledgments

Tom Decroos is supported by the Research Foundation-Flanders (FWO-Vlaanderen). Jesse Davis is partially supported by KU Leuven Research Fund (C14/17/07, C32/17/036) and Research Foundation - Flanders (EOS No. 30992574, G0D8819N). The authors thank Wyscout for supplying the event stream data used in this extended abstract.

⁴<https://www.skysports.com/football/news/11661/11540002/long-shots-work-andre-schurrle-and-ruben-neves-buck-the-trend>

References

- [Altman, 2015] Daniel Altman. Beyond Shots: A New Approach to Quantifying Scoring Opportunities. OptaPro Analytics Forum, 2015.
- [Beal *et al.*, 2020] Ryan Beal, Narayan Changder, Timothy J Norman, and Sarvapali D Ramchurn. Learning The Value of Teamwork to Form Efficient Teams. In *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence*, 2020.
- [Bransen *et al.*, 2019] Lotte Bransen, Pieter Robberechts, Jan Van Haaren, and Jesse Davis. Choke or Shine? Quantifying Soccer Players' Abilities to Perform Under Mental Pressure. In *MIT Sloan Sports Analytics Conference*, 2019.
- [Caley, 2015] Michael Caley. Premier League Projections and New Expected Goals. Cartilage Free Captain, 2015.
- [Chen and Guestrin, 2016] Tianqi Chen and Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785–794. ACM, 2016.
- [Decroos and Davis, 2019] Tom Decroos and Jesse Davis. Player Vectors: Characterizing Soccer Players' Playing Style from Match Event Streams. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 2019.
- [Decroos *et al.*, 2018] Tom Decroos, Jan Van Haaren, and Jesse Davis. Automatic Discovery of Tactics in Spatio-Temporal Soccer Match Data. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 2018.
- [Decroos *et al.*, 2019] Tom Decroos, Lotte Bransen, Jan Van Haaren, and Jesse Davis. Actions Speak Louder Than Goals: Valuing Player Actions in Soccer. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1851–1861. ACM, 2019.
- [Fernández *et al.*, 2019] Javier Fernández, Luke Bornn, and Dan Cervone. Decomposing the Immeasurable Sport: A Deep Learning Expected Possession Value Framework for Soccer. In *MIT Sloan Sports Analytics Conference*, 2019.
- [Jaspers *et al.*, 2018] Arne Jaspers, Tim Op De Beéck, Michel S Brink, Wouter GP Frencken, Filip Staes, Jesse J Davis, and Werner F Helsen. Relationships between the External and Internal Training Load in Professional Soccer: What Can We Learn from Machine Learning? *International Journal of Sports Physiology and Performance*, 13(5):625–630, 2018.
- [Kim *et al.*, 2018] Wonjun Kim, Sung-Won Moon, Jiwon Lee, Do-Won Nam, and Chanho Jung. Multiple Player Tracking in Soccer Videos: An Adaptive Multiscale Sampling Approach. *Multimedia Systems*, 24(6):611–623, 2018.
- [Liu and Schulte, 2018] Guiliang Liu and Oliver Schulte. Deep Reinforcement Learning in Ice Hockey for Context-Aware Player Evaluation. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, pages 3442–3448, 2018.
- [Lucey *et al.*, 2014] Patrick Lucey, Alina Bialkowski, Mathew Monfort, Peter Carr, and Iain Matthews. Quality vs. Quantity: Improved Shot Prediction in Soccer Using Strategic Features from Spatiotemporal Data. In *MIT Sloan Sports Analytics Conference*, 2014.
- [Mackay, 2017] Nils Mackay. Predicting Goal Probabilities for Possessions in Football. Master's thesis, Vrije Universiteit Amsterdam, 2017.
- [Pappalardo *et al.*, 2019] Luca Pappalardo, Paolo Cintia, Paolo Ferragina, Emanuele Massucco, Dino Pedreschi, and Fosca Giannotti. PlayeRank: Data-driven Performance Evaluation and Player Ranking in Soccer via a Machine Learning Approach. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(5):1–27, 2019.
- [Prokhorenkova *et al.*, 2018] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. CatBoost: Unbiased Boosting with Categorical Features. In *Advances in Neural Information Processing Systems*, pages 6639–6649, 2018.
- [Robberechts *et al.*, 2019] Pieter Robberechts, Jan Van Haaren, and Jesse Davis. Who Will Win It? An In-game Win Probability Model for Football. *arXiv preprint arXiv:1906.05029*, 2019.
- [Routley and Schulte, 2015] Kurt Routley and Oliver Schulte. A Markov Game Model for Valuing Player Actions in Ice Hockey. In *Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence*, pages 782–791, 2015.
- [Rudd, 2011] Sarah Rudd. A Framework for Tactical Analysis and Individual Offensive Production Assessment in Soccer Using Markov Chains. In *New England Symposium on Statistics in Sports*, 2011.
- [Schulte *et al.*, 2017] Oliver Schulte, Mahmoud Khademi, Sajjad Gholami, Zeyu Zhao, Mehrsan Javan, and Philippe Desaulniers. A Markov Game Model for Valuing Actions, Locations, and Team Performance in Ice Hockey. *Data Mining and Knowledge Discovery*, 31(6):1735–1757, 2017.
- [Shitrit *et al.*, 2011] Horesh Ben Shitrit, Jerome Berclaz, Francois Fleuret, and Pascal Fua. Tracking Multiple People under Global Appearance Constraints. In *2011 International Conference on Computer Vision*, pages 137–144. IEEE, 2011.
- [Singh, 2019] Karun Singh. Introducing Expected Threat. <https://karun.in/blog/expected-threat.html>, 2019. Accessed: 2019-06-21.
- [Van Roy *et al.*, 2020] Maaïke Van Roy, Pieter Robberechts, Tom Decroos, and Jesse Davis. Valuing On-the-Ball Actions in Soccer: A Critical Comparison of xT and VAEP. In *AAAI 2020 Workshop on Artificial Intelligence in Team Sports*, 2020.