

# Measuring Player Performance Through Predicting Team Actions

But Could They Do It On A Cold, Rainy Night in Stoke?

Olly Craven

## 1 Introduction

Models to value actions within football have existed for over a decade and many people have created different versions within this space, all making different design decisions. The football analytics community has developed many different models to try and value the contribution of players to their contracted club.

From a club perspective, this has been pursued in order to maximise the footballing output of their financial outlay by finding players who are cheaper than they should be or who will accept lower wages.

This paper details further work on a novel machine learning framework for forecasting football results building from fundamental actions upwards to predict games and seasons.

A specifically designed model will learn the impact of the players on the pitch on both action choice and outcomes. This allows a team to play differently with a change in personnel as well as isolating the impact of specific players or coaches on the overall team style. It provides a better attempt at modelling how replacing a player would impact how the team plays compared to current models, which primarily attempt to adjust for the tactical style of the player or team after the fact.

## **2 Aims**

There are a number of areas of player valuation that have proven difficult to advance in a meaningful and rigorous way. This research builds on previous work by Ben Torvaney<sup>[1]</sup> at Statsbomb Conference 2022.

The vast majority of clubs and adjacent organisations do not have the ability to collect or pay for highly detailed tracking data or similar and must make do with event-level data. Therefore new ways of deriving context-based metrics without being able to see the full context of each action are necessary. Alongside the issue of access to top-level data, there are also specific methodological issues that have plagued many research projects in this area.

### **2.1 Defensive Work Valuation**

It is very difficult to value defensive work, given that in event-level data only direct involvement is captured, whether that is by tackling, intercepting or pressing the ball carrier. The valuable work of marking a player so effectively that they are never a good option is very difficult to value from event level data. Kullowatz<sup>[2]</sup> created an extension of his G+ model in an attempt to value work done that was not connected to an event. He did this by creating zones that a player was proportionally responsible for, then added the G+ that a player allowed through their zone. The issue with this method is that it is not always clear cut who is responsible for what, and Kullowatz's method is still an unrobust way of measuring defensive contribution.

### **2.2 Off-Ball Valuation**

The EPV framework<sup>[3]</sup> provides an attempt at valuing purely off-ball contribution. The model significantly credits players for being in a good position to receive the ball in space, for example. It does continue to struggle with other aspects of off-ball contribution though. Even with the help of tracking data, the use of dummy runs and similar to create space has proved difficult to value as it is hard to divide credit between those who created the space and those who utilise it. The core problem therefore remains that it is remarkably difficult to capture the value of interactions between players when they have nothing directly to do with the actions happening on the ball.

### 3 Methodology

The dataset consists of the full free-released 15-16 Premier League season as well as all of the matches from half of the teams in the 16-17 and 17-18 seasons from Statsbomb.

The first step was to extract the lineups and coaches of the attacking and defending teams as well as the game state information, including location and which team is at home, for each action in the dataset.

A purposely designed neural network, shown in Figure 1 below, then learnt to predict which type of action would likely come next: successful pass, failed pass resulting in turnover, successful ball carry, failed ball carry resulting in turnover or a shot. It would also, for each of these actions, predict the parameters of a distribution that described the expected end location of that type of action for a given situation. These outputs collectively could then be used to simulate games between lineups and determine an estimate of the probability of each game outcome.

The model learns three probabilistic representations inspired by variational autoencoders, one each of the attacking team, the defending team and the game state. These representations turn the long string of binary inputs of which players are on the pitch into a much smaller set of inputs without losing information. Along with the probabilistic nature of the outputs, this should limit the problems of overfitting, though this cannot be properly investigated as there is not enough data to both train the model and provide a reasonable out of sample test dataset as it would require significant assumptions about how to input dozens of players from out-of-league transfers and team promotions.

For each non-shot action type, the model learns the parameters of a two-dimensional multivariate normal distribution. These parameters are the mean of each dimension and the three parameters that make up the lower-triangular decomposition of the covariance matrix. This method was chosen as it was well handled by the library used for the model as well as to speed up the random sampling of the distributions. To draw from a two-dimensional multivariate normal distribution:

$x$  is a vector randomly sampled from the 2D distribution with covariance matrix  $\Sigma$

$u$  is the vector of the means of each dimension of  $x$

$z$  is a vector of random variables drawn from the standard normal

$M$  is a lower triangular matrix such that:

$$M \cdot M^T = \Sigma$$

And therefore:

$$x = u + Mz$$

Having the model output the lower triangular matrix  $M$  directly significantly sped up the simulations as it removed the requirement to decompose the covariance matrix each time.

Once the model was trained, it was used to simulate the 2017-2018 Premier League season, the last one in its dataset. This season was chosen as it was able to make the most of the available data while minimising the assumptions needed by eliminating out-of-sample inputs. As this model could be considered generative due to its layout, measuring it by its ability to recreate data it has already seen is useful, but not definitive.

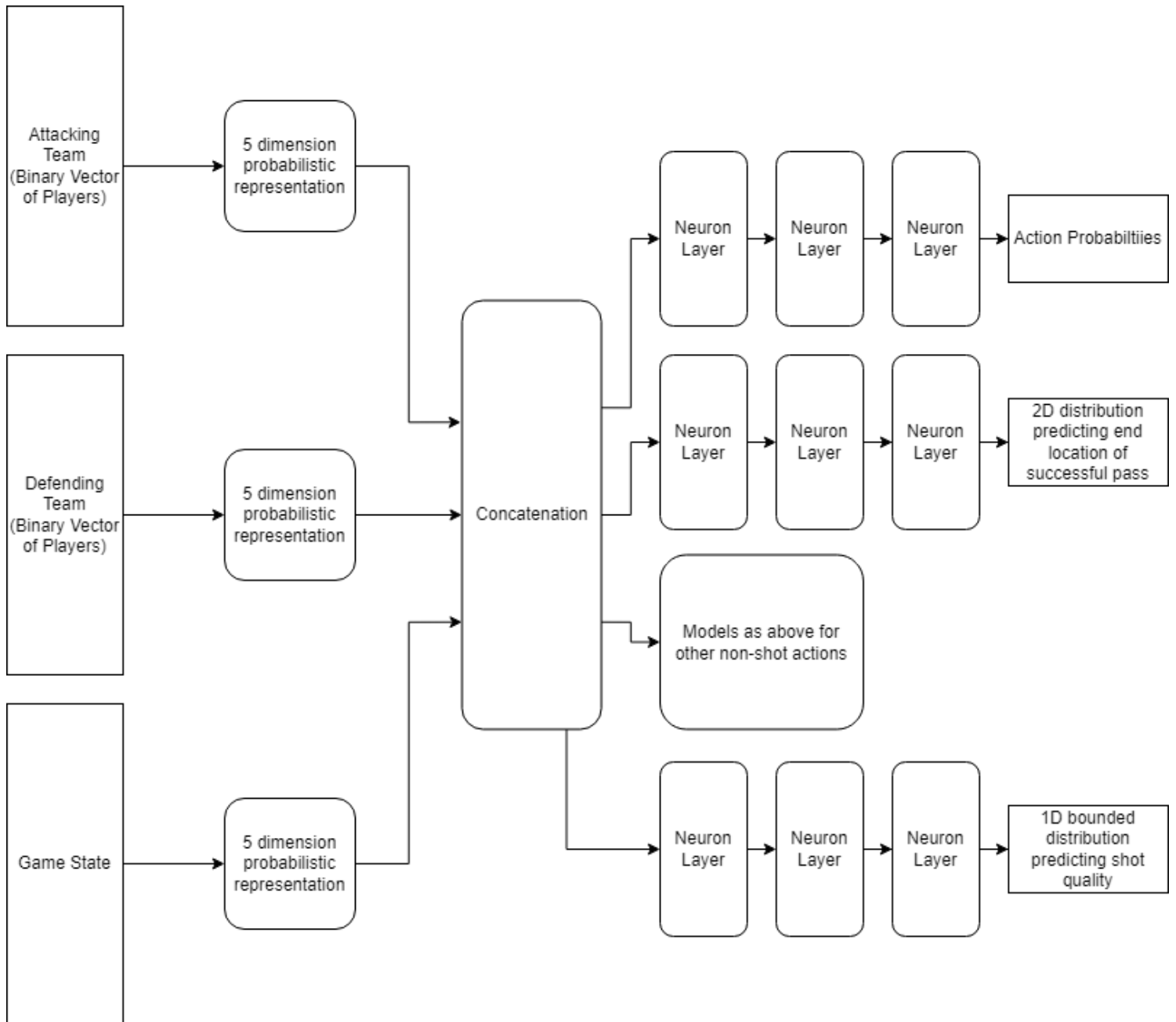
There are three methods of validation that were considered during the methodological process.

First, the appropriate inputs for the model were investigated. While these had to be changed due to discovered constraints, the inputs chosen for the final model still represent most of the meaningful inputs in theory.

The second is how well the model generated outputs are able to capture the features of the training data it was given. This will be explored further in section 4.

The final method of validation is testing the model on data it hasn't seen. Due to the nature of the dataset, this is very difficult to achieve. The model cannot make predictions for players it has never seen in its training data. The training data only covers three seasons of one league and given the quantity of transfers that occur it is very difficult to predict even the next season without considerably more available data.

Figure 1: Diagram showing the simplified layout of the proposed model.



## 4 Recreation Results

The first test of the model is whether it can replicate relatively realistic results from data that it has seen. The 2017-2018 Premier League season was used for this as it made use of the whole dataset without using information from future seasons. After many simulations with these lineups, the model provides reasonably ordered final tables and the final average table, shown in Table 1, is close to the actual result in points terms with a couple of notable exceptions. The standard deviation of points in a season is between 7 and 8 for each team, so the vast majority are comfortably within the margin of error we'd expect for a single season.

The core reasoning for the disparities is likely to be that the model has not fully learned the distribution of shot values. The model underestimates xG slightly, producing an average xG/shot of around 0.06 compared to the actual xG/shot average of around 0.10. This results in a smaller distribution of xGD and GD in the simulations compared to the actual season and therefore also a reduced distribution of points, as there will be more draws in lower scoring games. It therefore generally underestimates dominant teams and overestimates underdog teams.

The interesting overperformers though are Man Utd and Burnley. It is possible that the model struggles to pick up on the effectiveness of the defensive styles of Mourinho and Burnley, as they concede some of the lowest xG against per shot in the actual season. If the model was better at predicting shot value, it may also perform better at identifying when teams defend low and prevent high value shots.

The interesting underperformer is Southampton. The model is likely picking out that Southampton performed very well with a very similar set of players the previous seasons under Koeman and Puel and did not predict they would fall off as badly as they did.

On a match level, the model does provide some useful information. The Brier Skill Score (BSS) is a measure of how much better a model performs than a naive baseline prediction (based on the probability that the home team wins or draws) where 0 is as good as the baseline and 1 is when every probability prediction is 0 or 1 and correct in every instance. For this model, the BSS is 0.084 which competes with established models. For example, the match level predictions made by FiveThirtyEight (Boice and Wezerek, 2023) have an overall BSS of 0.067 and a BSS of 0.125 for the specific season simulated.

Table 1: Table showing the average goal difference and points for each team over the simulations compared to the real result of the season.

Team	Avg Simulated xG Difference	Avg Simulated Goal Difference	Avg Simulated Points	Real xG Difference	Real Goal Difference	Real Points
Man City	38.9	37.7	76.2	54.8	79	100
Arsenal	27.5	27.2	69.6	20.5	23	63
Tottenham	24.4	25.3	68.6	30.8	38	77
Liverpool	22.2	21.9	67.7	39.1	46	75
Chelsea	19.8	19.6	65.1	20.6	24	70
Man Utd	18.2	18.2	64.3	15.0	40	81
Southampton	2.9	3.9	55.6	-5.3	-19	36
Bournemouth	-2.1	-1.5	51.8	-20.4	-16	44
Watford	-2.8	-3.4	49.9	-6.8	-20	41
West Ham	-7.4	-7.4	47.2	-18.4	-20	42
Leicester	-8.8	-9.4	46.6	1.9	-4	47
Brighton	-9.2	-9.6	46.5	-13.8	-20	40
Everton	-10.4	-10.3	45.7	-11.9	-14	49
Stoke	-13.6	-13.8	44.3	-26.0	-33	33
Swansea	-16.2	-14.6	43.4	-27.0	-28	33
Crystal Palace	-12.6	-14.1	43.4	5.3	-10	44
Huddersfield	-14.8	-15.5	42.7	-17.2	-30	37
Newcastle	-17.2	-16.5	42.5	-10.8	-8	44
Burnley	-17.4	-16.7	42.1	-18.9	-3	54
West Brom	-21.5	-21.0	39.5	-11.4	-25	31

The lower level statistics generated by the model can also be analysed. In order to summarise the ability of the model across these statistics, the Pearson correlation coefficient is calculated for each season between the predicted statistics and the real statistics (Table 2). The average and standard deviation of these coefficients is then found for each statistic across all of the simulation instances, as is the correlation between the overall average of each statistic for each team over all simulations and the real statistics (Table 3).

The model provides strong and consistent correlation between the simulations and the actual season across passing statistics, but it struggles with the quality of shots (xG/Shot). There is some ability for the model to predict xG/Shot for a given team but is essentially unable to provide a useful prediction of xG/Shot against.

Table 2: Table showing the mean and standard deviation of the Pearson Correlation Coefficients of each seasons simulated statistics against the real statistics.

<b>Statistic</b>	<b>Pearson Average</b>	<b>Pearson StDev</b>
Points	0.72	0.08
Goal Difference	0.77	0.06
xGF	0.79	0.06
xGA	0.70	0.08
xGD	0.84	0.04
Shots For	0.71	0.09
Shots Against	0.82	0.04
xG/Shot For	0.31	0.13
xG/Shot Against	-0.05	0.21
Possession	0.96	0.01
Pass Success For	0.97	0.01
Pass Success Against	0.83	0.04



Table 3: Table showing the Pearson Correlation Coefficients of the average simulated statistics across all seasons against the real statistics

<b>Statistic</b>	<b>Pearson Average</b>
Points	0.881
Goal Difference	0.896
xGF	0.873
xGA	0.785
xGD	0.898
Shots For	0.872
Shots Against	0.845
xG/Shot For	0.366
xG/Shot Against	-0.113
Possession	0.962
Pass Success For	0.973
Pass Success Against	0.868

## 5 Manager Swapping

An interesting use-case of this model is estimating the impact that a different manager may have on a given team. To show this capability, simulations have also been run for scenarios in which Pep Guardiola is swapped with another Premier League manager from that season. The difference in the average points received between Man City with Pep and with the replacement and the difference in the average points between the replacement manager's original team and when that team is managed by Pep are shown in Table 4.

The model does clearly differentiate between the qualities of different managers and how they fit with their teams and does not particularly correlate with team quality. Despite the model's prediction that Liverpool would come fourth in the 17-18 season, it also thinks that Jürgen Klopp would have the best chance of performing as well as Pep did with Man City, while Pep provides the lowest improvement when he replaces Klopp at Liverpool compared to him replacing the other managers.

While Mourinho and Conte would both comparatively struggle with City according to the model, Pep does significantly better with the 17-18 Chelsea squad than he does with the 17-18 Man Utd squad.

Table 4: The average impact of swapping managers compared to the simulated season with correct team managers.

<b>Simulated Manager of Man City</b>	<b>Simulated Team that Pep is Managing</b>	<b>Man City Points Difference vs Simulated Actual</b>	<b>Pep's Team Points Difference vs Simulated Actual</b>
Pochettino	Tottenham	-4.2	4.5
Mourinho	Man Utd	-4.9	5.8
Wenger	Arsenal	-2.4	3.3
Conte	Chelsea	-5.4	8.1
Klopp	Liverpool	-2.3	2.0
Dyche	Burnley	-9.7	4.8

## 6 Player Transfer

The initial test of the model is one in which there should be a fairly obvious change in quality. For this, the hypothetical swap transfer of Kane and Rondon was used. The model provided a reasonably significant difference in the simulated results. Table 5 shows the results for both teams in the two simulations. A £2 million difference in prize money is roughly equivalent to ending a position higher on average at the lower end of the table, but at the top end of the table it is less clear, as qualification for Europe also makes a significant difference.

Given the big change in quality between the hypothetical transfer, it does make a considerable difference to the outcomes that both teams could expect. It would rule Tottenham out of any hope of winning the league, and cost them £7 million in prize money in the average season. It cuts West Brom's chances of being relegated by 4%, and nearly £4 million in extra money. This is smaller than the effect on Tottenham, but likely reflects the lack of quality in West Brom's lineup in terms of getting the ball to Kane in good situations in the first place or in providing quality runners to use his playmaking.

Table 6 demonstrates that the model is also capable of noticing changes in quality in traditionally difficult to measure positions like central and defensive midfield. Swapping Man City's Fernandinho and Stoke's Joe Allen does result in a significant difference in end of season results.

However, the model does struggle both with players who are of much more similar quality and with the areas of play which are difficult to measure. The two examples of this that were investigated were the Sanchez-Mkhitaryan swap deal and the Van Dijk transfer from Southampton to Liverpool. The results suggested that all teams involved would be doing better *without* the transfers occurring. Given that both of these transfers should be reasonably clear-cut, I've therefore concluded that there is an error in the method.

Further research would likely shed light on this issue. I think it is likely that the representations of the lineups are the issue, possibly that the representation is actually not small enough to properly generalise. It seems that currently any transfer that is not already clear cut in terms of quality causes issues, except with respect to managers.

Ignoring the latter results for a moment, the wider methodology of comparing the outcome of matches with and without a certain change is useful. The model keeps track of where each team finishes in each simulated season. This can then be combined with the financial benefits of both Premier League finishing position and from qualifying for Europe. Once the model's issues have been ironed out, it is likely to be useful in predicting the estimated financial result of a transfer or injury for those changes it currently struggles with.

Table 5: The results of the simulations for both teams when Kane and Rondon are swapped and play in each others' games.

Description	Average Goal Difference in Simulations	Average Points in Simulations	Probability of Winning the Premier League	Probability of Being Relegated	Average Prize Money from PL and European Qual (£mn)
Tottenham with Kane	25.3	68.6	0.13	0.00	121.16
Tottenham with Rondon	14.9	62.0	0.00	0.00	114.37
West Brom with Rondon	-21.0	39.5	0.00	0.42	89.89
West Brom with Kane	-14.7	43.5	0.00	0.38	93.62

Table 6: The results of the simulations for both teams when Fernandinho and Allen are swapped and play in each others' games.

Description	Average Goal Difference in Simulations	Average Points in Simulations	Probability of Winning the Premier League	Probability of Being Relegated	Average Prize Money from PL and European Qual (£mn)
Man City with Fernandinho	37.7	76.2	0.47	0.00	127.48
Man City with Allen	32.2	72.0	0.33	0.00	123.85
Stoke with Allen	-13.8	44.3	0.00	0.19	93.83
Stoke with Fernandinho	-7.6	47.8	0.00	0.18	97.23

## 7 Tactical Insight

The model can be used to output 5-dimension representations that it has learnt directly from the actions of the attacking and defending lineups. These could provide useful insight into the approach of each team but are difficult to interpret easily. For this reason, a similarity matrix has been created for attacking representations and for defending representations.

Figure 2 shows the similarity score between the two representations that the model learns for each team. A score of 1 suggests that the teams are essentially identical in approach, while a score of 0 suggests that they are the most different of any two teams, which is Man City and Leicester on the attacking side and Tottenham and West Brom on the defending side.

As we can see, the model considers Pep Guardiola's Man City to be significantly different to almost all other teams, both on attack and defence. Pochettino's Spurs are also very distinct on the defensive side.

Figure 2: The similarity score between the representations of each team's most common lineup separated into attacking and defending sides.

Attacking	Arsenal	Bournemouth	Brighton	Burnley	Chelsea	Crystal Palace	Everton	Huddersfield	Leicester	Liverpool	Man City	Man Utd	Newcastle	Southampton	Stoke	Swansea	Tottenham	Watford	West Brom	West Ham
Arsenal	1.00	0.57	0.41	0.15	0.73	0.34	0.18	0.28	0.14	0.69	0.77	0.76	0.26	0.52	0.32	0.41	0.75	0.34	0.25	0.48
Bournemouth	0.57	1.00	0.66	0.56	0.63	0.71	0.44	0.63	0.48	0.56	0.40	0.66	0.62	0.74	0.67	0.70	0.72	0.73	0.61	0.74
Brighton	0.41	0.66	1.00	0.63	0.53	0.74	0.59	0.77	0.63	0.49	0.29	0.45	0.68	0.80	0.81	0.87	0.57	0.72	0.77	0.83
Burnley	0.15	0.56	0.63	1.00	0.26	0.72	0.52	0.72	0.67	0.22	0.61	0.25	0.71	0.53	0.74	0.64	0.33	0.73	0.79	0.61
Chelsea	0.73	0.62	0.53	0.26	1.00	0.48	0.38	0.45	0.26	0.88	0.68	0.72	0.45	0.62	0.45	0.57	0.86	0.46	0.40	0.57
Crystal Palace	0.34	0.71	0.74	0.74	0.48	1.00	0.68	0.83	0.70	0.47	0.19	0.46	0.82	0.75	0.83	0.80	0.55	0.96	0.81	0.74
Everton	0.16	0.44	0.59	0.72	0.38	0.68	1.00	0.74	0.62	0.39	0.07	0.27	0.68	0.57	0.64	0.68	0.39	0.65	0.69	0.55
Huddersfield	0.28	0.63	0.77	0.72	0.45	0.83	0.74	1.00	0.75	0.43	0.17	0.36	0.76	0.70	0.79	0.84	0.48	0.81	0.87	0.74
Leicester	0.14	0.46	0.63	0.67	0.26	0.70	0.62	0.75	1.00	0.26	0.00	0.21	0.53	0.60	0.64	0.63	0.30	0.69	0.69	0.60
Liverpool	0.69	0.58	0.49	0.22	0.88	0.47	0.39	0.43	0.26	1.00	0.53	0.71	0.40	0.62	0.42	0.54	0.80	0.45	0.37	0.53
Man City	0.77	0.40	0.29	0.01	0.68	0.19	0.07	0.17	0.00	0.63	1.00	0.58	0.14	0.37	0.18	0.30	0.62	0.18	0.13	0.35
Man Utd	0.76	0.66	0.45	0.25	0.72	0.46	0.27	0.36	0.21	0.71	0.55	1.00	0.41	0.57	0.43	0.47	0.83	0.46	0.34	0.51
Newcastle	0.26	0.62	0.66	0.71	0.43	0.82	0.68	0.76	0.58	0.40	0.14	0.41	1.00	0.82	0.84	0.75	0.50	0.79	0.83	0.66
Southampton	0.52	0.74	0.80	0.53	0.55	0.75	0.57	0.70	0.60	0.62	0.37	0.57	0.62	1.00	0.71	0.79	0.68	0.73	0.65	0.83
Stoke	0.32	0.67	0.81	0.74	0.45	0.83	0.64	0.79	0.54	0.42	0.18	0.43	0.84	0.71	1.00	0.82	0.53	0.81	0.86	0.78
Swansea	0.41	0.70	0.87	0.64	0.57	0.80	0.68	0.84	0.63	0.54	0.30	0.47	0.75	0.79	0.82	1.00	0.61	0.77	0.81	0.88
Tottenham	0.75	0.72	0.57	0.33	0.86	0.55	0.39	0.48	0.30	0.80	0.62	0.63	0.50	0.73	0.88	0.53	0.61	1.00	0.54	0.63
Watford	0.34	0.73	0.72	0.73	0.46	0.96	0.65	0.81	0.69	0.45	0.18	0.46	0.79	0.73	0.81	0.77	0.54	0.54	1.00	0.79
West Brom	0.25	0.61	0.77	0.79	0.40	0.81	0.69	0.87	0.69	0.37	0.13	0.34	0.83	0.85	0.86	0.81	0.45	0.45	0.79	1.00
West Ham	0.48	0.74	0.93	0.61	0.57	0.74	0.55	0.74	0.60	0.53	0.35	0.51	0.65	0.83	0.78	0.86	0.63	0.72	1.00	0.73
Defending	Arsenal	Bournemouth	Brighton	Burnley	Chelsea	Crystal Palace	Everton	Huddersfield	Leicester	Liverpool	Man City	Man Utd	Newcastle	Southampton	Stoke	Swansea	Tottenham	Watford	West Brom	West Ham
Arsenal	1.00	0.84	0.57	0.58	0.82	0.53	0.72	0.53	0.66	0.66	0.52	0.67	0.57	0.88	0.51	0.67	0.51	0.75	0.41	0.65
Bournemouth	0.84	1.00	0.64	0.64	0.82	0.55	0.77	0.68	0.74	0.57	0.21	0.67	0.67	0.88	0.55	0.82	0.38	0.77	0.52	0.73
Brighton	0.57	0.64	1.00	0.76	0.60	0.69	0.67	0.75	0.72	0.51	0.20	0.64	0.62	0.55	0.63	0.73	0.21	0.74	0.75	0.57
Burnley	0.58	0.64	0.76	1.00	0.55	0.64	0.67	0.69	0.75	0.53	0.33	0.60	0.74	0.62	0.86	0.71	0.31	0.74	0.64	0.68
Chelsea	0.82	0.82	0.60	0.65	1.00	0.48	0.69	0.58	0.80	0.68	0.30	0.59	0.69	0.58	0.57	0.69	0.39	0.81	0.48	0.75
Crystal Palace	0.53	0.55	0.68	0.64	0.48	1.00	0.57	0.51	0.51	0.39	0.28	0.60	0.49	0.68	0.58	0.61	0.32	0.57	0.51	0.41
Everton	0.72	0.77	0.67	0.67	0.69	0.57	1.00	0.65	0.65	0.60	0.58	0.65	0.57	0.71	0.61	0.77	0.42	0.79	0.46	0.72
Huddersfield	0.53	0.68	0.75	0.69	0.58	0.57	0.65	1.00	0.70	0.37	0.05	0.56	0.70	0.55	0.58	0.84	0.14	0.64	0.77	0.62
Leicester	0.66	0.74	0.72	0.51	0.80	0.51	0.65	0.70	1.00	0.58	0.22	0.55	0.81	0.52	0.49	0.73	0.24	0.80	0.66	0.74
Liverpool	0.66	0.57	0.51	0.53	0.68	0.39	0.60	0.37	0.58	1.00	0.47	0.60	0.43	0.45	0.49	0.47	0.45	0.73	0.32	0.56
Man City	0.32	0.21	0.20	0.33	0.30	0.26	0.30	0.65	0.22	0.47	1.00	0.33	0.17	0.33	0.40	0.15	0.56	0.34	0.00	0.28
Man Utd	0.67	0.67	0.64	0.60	0.59	0.60	0.68	0.56	0.59	0.60	0.53	1.00	0.46	0.69	0.55	0.66	0.45	0.72	0.42	0.30
Newcastle	0.57	0.67	0.62	0.62	0.69	0.49	0.57	0.70	0.81	0.43	0.17	0.46	1.00	0.53	0.68	0.72	0.22	0.65	0.65	0.73
Southampton	0.68	0.68	0.55	0.62	0.58	0.68	0.71	0.55	0.52	0.45	0.33	0.69	0.53	1.00	0.59	0.68	0.53	0.60	0.39	0.57
Stoke	0.51	0.55	0.63	0.88	0.57	0.58	0.61	0.58	0.65	0.49	0.45	0.55	0.68	0.59	1.00	0.61	0.33	0.65	0.53	0.66
Swansea	0.67	0.82	0.73	0.71	0.69	0.61	0.77	0.84	0.73	0.47	0.15	0.66	0.72	0.88	0.61	1.00	0.28	0.72	0.64	0.71
Tottenham	0.51	0.38	0.21	0.31	0.39	0.32	0.42	0.14	0.24	0.45	0.33	0.45	0.22	0.53	0.33	0.28	1.00	0.37	0.08	0.33
Watford	0.75	0.77	0.74	0.74	0.81	0.39	0.42	0.64	0.80	0.73	0.54	0.72	0.55	0.80	0.65	0.83	0.28	1.00	0.56	0.73
West Brom	0.41	0.52	0.75	0.64	0.49	0.51	0.48	0.77	0.68	0.32	0.00	0.42	0.65	0.39	0.53	0.64	0.00	0.37	1.00	0.49
West Ham	0.65	0.73	0.57	0.68	0.75	0.41	0.72	0.62	0.74	0.56	0.26	0.58	0.73	0.57	0.66	0.71	0.33	0.73	0.56	1.00

## 8 Conclusions

While it does have a few issues, there is evidence to suggest that the initial framework is a useful way of thinking about valuing players and teams while avoiding the difficult methodological questions that plague models that try to value individual actions.

The framework provides an easy way of scaling directly from single actions to game level to season level predictions. It provides a method of accounting for contextual impact with respect to the game state and both teams on the pitch in a rigorous manner, unlike action valuing models which require significant assumptions and post-hoc adjustment. This allows a clearer way of estimating an individual player's impact on the teams results and provides a base from which the simulation can be made more complex to include more information without requiring a significant architectural change on the model level.

This framework does, however, require more data and more computational power than the alternatives compared to the action valuing models. With an action value model, most of the computational requirement is front-loaded onto the actual valuing model, then there is only a small overhead to estimate new actions or to manipulate the individual valued actions into collated information for use by decision-makers. This framework requires significant ongoing computational use as each new suggested input will require a large number of simulations to produce a reasonable result.

The model provides some very interesting results on the manager results, with clear differences between manager styles and how they fit with different squads. It does not just map team quality directly onto the manager, which is what could have been expected given the poor performance of the model with respect to player transfers. This part of the model output is probably the most useful in terms of actual decision making at this time.

It is disappointing that the model's ability to interpret transfers is not fully working as intended. It can measure the impact of clear cut cases of improvement or worsening in obvious positions, but struggles more with players of closer quality. It is likely the case that the learned representations are the issue and it is very possible that with further research this issue could be solved. It does provide some proof of its usefulness in situations where the change in quality is clear, so future work may very well extend this ability to the much closer cases.

Overall, the framework provides a useful starting point for future research in this area. It develops on Torvaney's work to solve a number of issues and provide a more useful output for decision-makers. It does have problems, but these do seem to be solvable rather than fundamental issues with the approach taken.

## 9 Further Work

The first strand of further work is clearly fixing the issue of improper representation learning. The model does not use the common method of representation learning, which can be described simply as training a model that has to squeeze information through a narrow gap and then recreate it. Instead of this, the representations are learnt using the lineups as an input, trained on the actual on-pitch actions. The model may be under-trained or the representations may be too wide to force the model to learn only key information.

The second strand of work is on the shot quality issue. The best way forward for this may be training the model on everything except shots, and then freezing that model and training an add-on model using only shots. While this would remove some of the possible information that could be learnt in the representations, it solves the issue of scarcity of shots and therefore reduces the need for training set weighting.

The simulation portion could also be further developed to involve ideas like player injuries and preferred selection inputs so that the simulations become more dynamic with respect to which players play in each fixture to become more realistic.



## References

- [1] B. Torvaney, "Quality as an Emergent Property: A Case Study with xThreat", at StatsBomb Conference, London, UK, September 2022. Available: <https://www.youtube.com/watch?v=oMdXp9BhxTs>
- [2] M. Kullowatz, "Introducing goals subtracted: Where you aren't but you oughta be," American Soccer Analysis, 10-Aug-2022. [Online]. Available: <https://www.americansocceranalysis.com/home/2022/8/8/introducing-goals-subtracted-g->
- [3] J. Fernandez, L. Bornn, and D. Cervone, "Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer," MIT Sloan Sports Analytics Conference, 22-Feb-2019. [Online]. Available: <https://www.sloansportsconference.com/research-papers/decomposing-the-immeasurable-sport-a-deep-learning-expected-possession-value-framework-for-soccer>