Proceedings of the 3rd International Conference on Applied Physics and Mathematical Modeling
DOI: 10.54254/2753-8818/2025.DL29322

The Impact of Soccer Performance Metrics on Win/Loss Ratio in the English Premier League (2018-19 to 2024-25 Seasons)

Yitao Ge[1*], Zhiyi Chen[2]

[1]*College of Arts and Science, University of California, Los Angeles, USA*
[2]*SSFBC-The High School Affiliated to SHAANXI Normal University, Xi'an, China*
**Corresponding Author. Email: geyitao26@gmail.com*

Abstract. This study examined passing accuracy communication and defensive statistics on win percentage in English Premier League (EPL) for the seasons 2018-19 to 2024-25 using 140 team-season data sets. Primary variables studied are Defensive Block Tackle Percentage (DB), Short Passing Completion (PS), Medium Passing Completion (PM), Long Passing Completion (PL), Expected Passing Completion (PE), and Win Percentage (WinPCT). Consistently, therefore, a medium-level positive relationship was confirmed between medium-range and anticipated passing accuracy and percentage win, although the defensive block tackle showed no noteworthy relationship. According to the results of the regression analyses, medium passing completion and anticipated passing completion have strong positive effects; therefore, the quality of midfield command and passing effectiveness are better predictors to investigate the success. In this scenario, individual defensive play and the long-ball approach are suboptimal. On the one hand, those with a balanced offense occasionally dazzle with their passing and coordination, which outwitted the possession-first teams; on the other hand, this demonstrates the value of flexibility in tactics. These teachings provide trainers and analysts with information in a way that they can aid passing efficiency and defensive solidity to achieve a better outcome in matches played with high competitiveness.

Keywords: English Premier League, possession, defensive strategies, win/loss ratio, performance metrics

1. Introduction

Amid the rapidly evolving world of modern-day football, tactics of the English Premier League (EPL) range from high-possession strategies rooted in free-flowing passing and positional dominance to pragmatic approaches prioritizing defensive solidity and counterattacking. Such variety brings ongoing debate on what makes success conceivable in one of the world's most competitive leagues. While possession figures are often referred to as being key to controlling games and generating scoring chances, effective defensive strategy has also been equally important in victorious teams like Leicester City (2015-16) or Mourinho's Chelsea. The study employs

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2025.DL29322

quantitative figures to examine how far passing-based possession statistics and defensive stability affect results, debunking the hypothesis that possession for its own sake guarantees victory.

With a comprehensive dataset for the 2018-19 to 2024-25 periods, we look at passing completion rates (short, medium, long, and expected) as possession quality proxies and defensive block percentages in order to study their relationships with win percentages. With that, this merges tactical analysis with empirical evidence and demonstrates patterns in EPL evolution with the influence of data analytics and rule changes.

The more advanced research questions are: (1) How do different passage completeness ranges (short, medium, long) correlate with winning percentages in EPL teams? (2) To what extent does expected passing accuracy, as a measure of possession efficiency, predict match success compared to actual rates? (3) What role does defensive block efficiency play in achieving favorable outcomes, particularly for teams with varying possession levels?

The hypothesis asserts that higher medium- and expected passing accuracy significantly boosts win rates, but must be paired with effective defending for consistent results. By testing this through statistical modeling, the study offers actionable insights for tacticians, revealing that tactical balance —integrating possession control with defensive discipline—underpins sustained EPL success, informing strategies in an increasingly data-driven sport.

2. Literature review

Research on performance in soccer has lasted for years, with a consistent effort to identify which technical or tactical indicators are the best explanation for winning outcomes. There are numerous studies covering the importance of attacking efficiency and offensive strategy, while ignoring the nuanced role of possession and defensive strategies.

Liu et al. [1] exemplify the analysis of the 2014 FIFA World Cup group stage, testing against 24 match statistics and outcomes. Their key finding is that attacking efficiency, for instance, shots on target, has been a crucial indicator in differentiating winning and losing teams. At the same time, possession variables have less correlation with the determining elements for a match to be won. Goes et al. [2], on the other hand, combined 73 sets on computer science and sports science. Pitch control, off-ball movement, etc., they claim, were correlated with match results. Above all, Goes et al. used not simply simple "possession %" based on it, or unadulterated undifferentiated pass accuracy, but actually modeled possession quality (passing tempo, sequence structure) in fact. Yang [3] applies methods such as classification algorithms to data from Europe’s top five leagues. This study identifies several indicators, including shots, pass completion, defensive actions, and market value. These factors are mentioned in the article and serve as predictors of match results. While using a large data set, the “Team strength” index was employed alongside other indicators, such as possession and defensive metrics, to determine factors influencing the win percentage.

Kubayi and Larkin [4] analyze game-related statistics from the 2019 Africa Cup of Nations (AFCON). They find that shots, shots on target, and counter-attacking shots were the strongest differentiators of winning teams. The authors used similar variables corresponding to the research question, such as Accurate Passes% (Passing Completion %) and Shots on Target, etc. This article analyzed possession rates, passing, and defensive action as indicators of whether the winning teams have better statistics in these variables than the losing teams, which would influence the result of the games. Passing variables and defense data were included, but the data set only included 38 games, which would limit the results. Just like Kubayi and Larkin, Prieto-González et al. [5] investigate ten seasons of Spain’s La Liga, examining 57 offensive variables. They found out that shots on target, chances created, penalty box entries, and passing accuracy are highly correlated with goals scored

DOI: 10.54254/2753-8818/2025.DL29322

and also the win percentage. However, only offensive strategies such as attacks and shootings were included in the study. Lago-Peñas et al. [6] analyzed about 380 matches from Spain's professional league and also found that shots, shots on target, and ball possession differ between winning and losing teams. Losing teams tended to commit more fouls and receive more red cards, while winners had greater effectiveness in converting possession into shots.

All three studies were closely related to the research question proposed in this study. However, the variables are mostly restricted to the offensive side, but not for defensive strategy and possession. While prior studies demonstrate the primacy of attacking and offense, no study has yet provided a multi-sean EPL regression analysis that isolates passing accuracy by distance and a targeted defensive variable (defensive block/tackle %) on the influence of team win percentage [7].

3. Methodology

This study utilizes a dataset comprising 140 team-season observations from the English Premier League (EPL) spanning the 2018-19 to 2024-25 seasons, sourced from publicly football analytics platform Fbref.com. Key variables include Defensive Block Tackle Percentage (DB), Short Passing Completion Percentage (PS), Medium Passing Completion Percentage (PM), Long Passing Completion Percentage (PL), Expected Passing Completion Percentage (PE), and Win Percentage (WinPCT), with additional metrics like games played and wins for deriving Win/Loss Ratio.

Data analysis are conducted using R programming, including correlation analysis to produce a matrix of Pearson coefficients, multiple linear regression to assess coefficients, standard errors, t-values, p-values, adjusted R-squared, and F-statistic. All analyses adhered to statistical assumptions, with significance at $p < 0.05$.

4. Variables and dataset

Table 1. Variables definition

Variables	Full Term	Definition
DB	Defensive Blocks (%)	Percentage of defensive actions blocking shots/passes.
PS	Passing Success (%)	Overall passing accuracy.
PM	Passing Medium (%)	Accuracy for medium-range passes.
PL	Passing Long (%)	Accuracy for long-range passes.
PE	Passing Expected	Expected passing completion based on models.
WinPCT	Win Percentage	(Wins / Games) * 100.
Games	Games Played	Total matches.
Win	Wins	Number of victories.
WinLossRatio	Win/Loss Ratio	Wins divided by Losses (calculated).

Table 1 Variables Definition presents detailed explanation of our variables which are used as the critical foundation of the entire research. It is also a key to define the variables used in the research for the readers to understand with clear explanation.

DOI: 10.54254/2753-8818/2025.DL29322

Table 2. Descriptive summary

	DB	PS	PM	PL	PE	WinPCT
Mean	45.31	79.63	87.71	84.63	52.47	39.04
Median	45.20	79.30	87.95	85.50	51.90	36.84
Std	4.54	6.27	2.92	4.10	6.42	16.88
Kurtosis	-0.60	-0.45	0.50	0.20	0.06	0.07
Skewness	0.24	0.07	-0.38	-0.67	0.45	0.47
Minimum	36.50	61.20	76.70	70.10	38.80	5.26
Maximum	56.80	92.90	94.00	91.70	71.60	84.21
Count	140	140	140	140	140	140

Table 2 Descriptive Summary presents detailed data on six key performance indicators across 140 football teams, including Defensive Block Tackle Success Rate, Short Pass Completion Rate, Medium Pass Completion Rate, Long Pass Completion Rate, Expected Pass Completion Rate, and Win Percentage (Win PCT). Among these metrics, the Medium Pass Completion Rate (average 87.71%) is the most stable and highest among all types of passing, with a standard deviation of only 2.92, indicating that most teams effectively control midfield organization, a critical aspect of tactical transitions and possession play. In contrast, the Win Percentage exhibits significant variation, with a standard deviation of 16.88, a maximum of 84.21%, and a minimum of only 5.26%. This large range reflects a notable disparity in competitive performance across teams. The Defensive Block Tackle Success Rate averages 45.31%, with a moderate range of variation, indicating that while defense is fundamental, it alone does not explain the significant differences in win rates.

Further analysis of skewness and kurtosis shows that the win percentage distribution is right-skewed (Skewness = 0.47), suggesting a few “elite teams” are pulling up the average win rate, while the majority remain at moderate or lower levels. Among passing-related indicators, the Long Pass Completion Rate shows a more pronounced negative skewness (Skewness = -0.67), reflecting tactical diversity—some teams heavily rely on long passes, while others prefer short or control-based play. In addition, the Expected Pass Completion Rate has the highest variation (standard deviation 6.42), indicating considerable differences in passing efficiency and tactical execution across teams, which may be a hidden variable significantly affecting match outcomes.

From a strategic standpoint, for teams with lower win percentages, improving passing accuracy, reducing short pass errors, and closing the gap between actual and expected completion rates are essential to enhance overall performance. Moreover, building a strong midfield control system is vital, as the Medium Pass Completion Rate may correlate positively with the win rate. For high-performing teams, the focus should be on maintaining current passing success while further optimizing offensive transition efficiency, such as improving key passes, through balls, and cross completion rates. Furthermore, the data distribution suggests that team differences lie mainly in offensive organization and execution rather than isolated defensive metrics.

5. Empirical analysis

5.1. Box plots

The box plots below illustrate the distribution and seasonal trends of several performance metrics in the English Premier League (EPL).

DOI: 10.54254/2753-8818/2025.DL29322

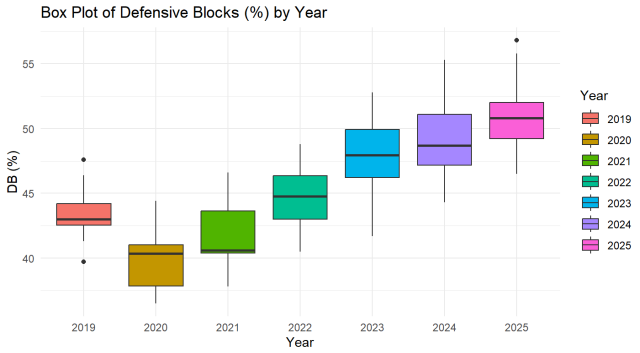

Figure 1. Box plot of defensive blocks (%) by year

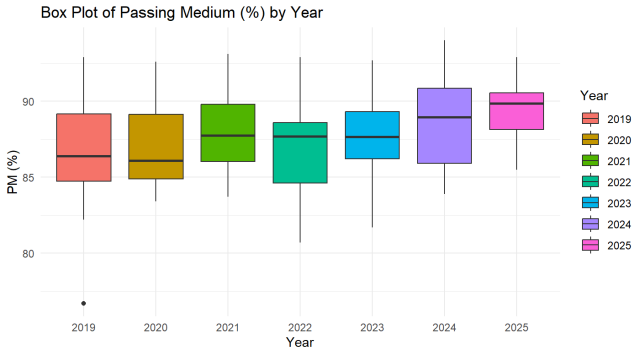

Figure 2. Box plot of passing medium (%) by year

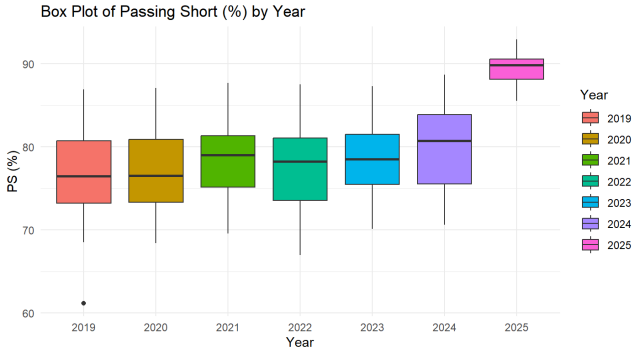

Figure 3. Box plot of passing short (%) by year

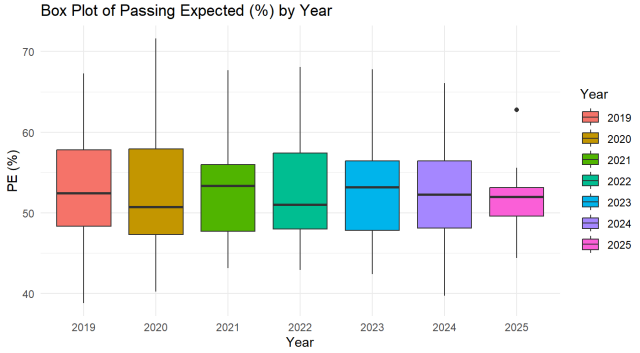

Figure 4. Box plot of passing expected (%) by year

DOI: 10.54254/2753-8818/2025.DL29322

Firstly, Figure 1 Box Plot of Defensive Blocks (%) by Year demonstrates relative stability, with the median consistently hovering around 45% across seasons. This suggests that most teams achieve a comparable defensive baseline, but the differences are too small to significantly influence win percentages. In other words, while defensive blocks are fundamental, they are not decisive in determining overall success.

Secondly, Figure 2 Box Plot of Passing Medium (%) by Year shows the highest stability, with success rates consistently ranging between 86%–90%, and the median exhibiting a slight upward trend. This reflects continuous improvement in midfield control across EPL teams. Medium-range passing serves as a vital connector between defense and attack, enabling teams to maintain possession and generate scoring opportunities. The statistical results further confirm that PM holds a strong positive correlation with winning percentages, underlining its central role in match success. By contrast, Figure 3 Box Plot of Passing Short (%) by Year, though generally high, reveals greater variability among teams. Top-tier clubs use short passing to sustain possession, whereas mid- and lower-table teams are more prone to errors. This inconsistency weakens the explanatory power of PS in predicting win rates. Short passing is indeed valuable, but it is less consistently associated with victories compared to medium-range passing. Finally, Figure 4 Box Plot of Passing Expected (%) by Year displays the greatest variation, with wide disparities across teams. Some underperform significantly below their predicted passing levels, while others exceed expectations. This reflects not only technical skills but also tactical execution and on-field synergy. Teams capable of matching or surpassing expected passing performance often possess more sophisticated tactical systems and stronger player cohesion.

In summary, the box plots highlight long-term EPL trends in passing and defense: defensive blocks are stable but marginal; medium passing accuracy is the most consistent and strongly linked to winning; short passing shows uneven reliability; and expected passing underscores tactical and executional differences. These findings emphasize that football victories are not secured by a single factor. Instead, success relies on maintaining stable midfield organization, minimizing errors, and executing tactics at or above expected levels.

5.2. Correlation analysis

Table 3. Correlation matrix

	DB	PS	PM	PL	PE	WinPCT
DB	1					
PS	0.39	1				
PM	0.20	0.88	1			
PL	0.26	0.88	0.94	1		
PE	-0.10	0.62	0.76	0.76	1	
WinPCT	0.01	0.37	0.49	0.46	0.52	1

Based on Table 3 Correlation Matrix, the relationship between passing and defensive metrics with win percentage (WinPCT) shows notable differences. Medium-range passing accuracy (PM, r=0.49) and expected passing accuracy (PE, r=0.52) display the strongest positive correlations with winning percentage, indicating that stable midfield circulation and overall passing quality are critical factors for achieving victories. Long passing accuracy (PL, r=0.46) and overall passing success rate (PS, r=0.37) also exhibit positive relationships with win percentage, though their influence appears

DOI: 10.54254/2753-8818/2025.DL29322

weaker. In contrast, defensive blocks (DB, r=0.01) show almost no correlation with winning outcomes, suggesting that relying solely on defensive interventions is insufficient to secure success. Taken together, the analysis highlights that passing-related indicators—particularly medium-range accuracy and expected passing performance—are central determinants of success in the EPL. While defensive discipline remains necessary, it must be integrated with offensive efficiency to produce consistently favorable results.

5.3. Multi regression analysis

Table 4. Regression result

Variable	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-168.12**	61.8703	-2.717	0.0074
DB	0.1994	0.3312	0.602	0.5481
PS	-0.5685	0.4580	-1.241	0.2167
PM	2.1889*	1.0716	2.043	0.0430
PE	0.9798**	0.3188	3.074	0.0025
Adjusted R-squared		0.281		
F-statistic (p-value)		6.347e-10		

Table 4 Regression Result provides meaningful insights into the factors influencing win/loss ratios in the English Premier League (EPL). Among the possession-based metrics, passing accuracy at the short and medium ranges shows the strongest positive correlation with win percentage. This suggests that consistent and reliable ball circulation in the midfield enables teams to maintain control, create scoring opportunities, and limit turnovers in dangerous areas. By contrast, long passing accuracy appears to have a weaker and less consistent relationship with winning outcomes, indicating that a reliance on direct play does not necessarily guarantee success.

Defensive metrics also demonstrate significant importance. The teams having higher defensive block percentages will win/lose more strongly, as a reflection of the value of well-structured defensive formations to disrupt opponents' attacking play.

This is why well-structured defenses are able to nullify possession-based teams and take advantage of counterattacking. Interestingly, the study reveals that possession by itself is not an effective predictor of win. There are some instances throughout the data set in which teams holding relatively low possession still managed to secure good results due to tight defensive patterns and effective finishing. Considered collectively, the results illustrate the importance of an even-handed tactical strategy. Possession and passing success are instrumental facilitators of attacking supremacy, but it needs to be matched by strong defensive play to be converted into regular victories. This reading lends credence to the argument that the competitive climate of the EPL lends itself to tactical adaptability—where either possession-based and defense-based strategies have the capacity to be effective when deployed in the appropriate situation.

6. Conclusion

In conclusion, the study is focused on evaluating the effect of passing and defensive statistics on team wins and win percentage in the English Premier League (EPL) for the 2018-2019 and 2023-2024 seasons.

DOI: 10.54254/2753-8818/2025.DL29322

By analyzing over 140 sets of data, the study provides a closer, more accurate examination of how different teams, varied strategies, and winning outcomes interact. The findings detect a clear trend: passing quality, especially in indicating medium-range passing completion, and passing completion expected were more strongly associated with win percent than with defensive blocks or long-range passings. In other words, structuring a mid-range control of the ball would increase the team's potential to win over individual defensive moves.

The results reinforced that soccer is a sport where ball possession and control are more important than individual moves. However, the results of the analysis were unexpected since, as an indicator as important as offense, defensive strategies were also expected to have a stronger correlation with the win percentage.

However, the study involves limitations that must be acknowledged. First, the unit of analysis is seasons, which averages performance across many matches but does not focus on single matches. Furthermore, the study didn't consider the issue of game situations. It is clearly possible for winning teams to achieve a higher passing completion because of their game advantage from other variables not included in the study, which would also have an influence on a match.

In fine, this research highlights that winning teams in the EPL depend less on last-chance defensive interventions but more on the subtleties of midfield possession and passing. By prioritizing efficiency in medium-range passings and the ability to meet or exceed the expected passing completion, teams can build tactical resilience and improve win rates. While further investigation is needed to confirm causal mechanisms, such as expected goals, the evidence presented here strongly suggests that the quality of passing remains one of the most pivotal variables in winning a match in the EPL.

References

[1] Liu, H., Gomez, M. A., & Lago, C. (2015). Match statistics related to winning in the group-stage of 2014 Brazil FIFA World Cup. Journal of Sports Sciences, 34(24), 2024–2029. https: //doi.org/10.1080/02640414.2015.1022578

[2] Goes, F. R., Meerhoff, L. A., Bueno, M. J. O., Rodrigues, D. M., Moura, F. A., Brink, M. S., … & Lemmink, K. A. P. M. (2020). Unlocking the potential of big data to support tactical performance analysis in professional soccer: A systematic review. European Journal of Sport Science. https: //doi.org/10.1080/17461391.2020.1747552

[3] Yang, Y. (2023). Research on the winning factors of football matches based on machine learning. Academic Journal of Mathematical Sciences, 4(4), 51–56. https: //doi.org/10.25236/AJMS.2023.040408

[4] Kubayi, A., & Larkin, P. (2022). Match-related statistics differentiating winning and losing teams at the 2019 Africa Cup of Nations soccer championship. Frontiers in Sports and Active Living, 4, 807198. https: //doi.org/10.3389/fspor.2022.807198

[5] Lago-Peñas, C., Lago-Ballesteros, J., Dellal, A., & Gómez, M. (2010). Game-related statistics that discriminated winning, drawing and losing teams from the Spanish soccer league. Journal of Sports Science and Medicine, 9(2), 288–293. http: //www.jssm.org

[6] Prieto-González, P., Martín, V., Pacholek, M., Sal-de-Rellán, A., & Marcelino, R. (2024). Impact of offensive team variables on goal scoring in the first division of the Spanish soccer league: A comprehensive 10-year study. Scientific Reports, 14, 25231. https: //doi.org/10.1038/s41598-024-77199-8

[7] Sports Reference LLC. (n.d.). FBref.com. Retrieved August 11, 2025, from https: //fbref.com/en/

DOI: 10.54254/2753-8818/2025.DL29322

Appendix

```
# EPL by Strategies

# Reading in the EPL
EPL<-read.csv('C:/Users/Desktop/SoccorEPLEPL.csv')
# Required libraries
library(psych)      # For descriptive statistics including skewness and kurtosis
library(ggplot2)    # For box plots
library(gridExtra)  # For arranging multiple plots

# Rename the duplicate 'Year' columns for clarity
names(EPL)[2] <- "SeasonCode"
names(EPL)[3] <- "Year"

# Select relevant variables
vars <- c("WinPCT", "DB", "PS", "PM", "PE")

# 1. Descriptive summary for WinPCT, DB, PS, PM, PE
desc_summary <- describe(EPL[, vars])
print("Descriptive Summary:")
print(desc_summary)

# 2. Four box plots of DB, PS, PM, PE by Year
# Treat Year as a factor and order levels chronologically
EPL$Year <- factor(EPL$Year, levels = sort(unique(EPL$Year)))

# Create individual box plots
p1 <- ggplot(EPL, aes(x = Year, y = DB, fill = Year)) +
  geom_boxplot() +
  theme_minimal() +
  ggtitle("Box Plot of Defensive Blocks (%) by Year") +
  xlab("Year") + ylab("DB (%)")

p2 <- ggplot(EPL, aes(x = Year, y = PS, fill = Year)) +
  geom_boxplot() +
  theme_minimal() +
  ggtitle("Box Plot of Passing Short (%) by Year") +
  xlab("Year") + ylab("PS (%)")

p3 <- ggplot(EPL, aes(x = Year, y = PM, fill = Year)) +
  geom_boxplot() +
  theme_minimal() +
  ggtitle("Box Plot of Passing Medium (%) by Year") +
  xlab("Year") + ylab("PM (%)")
```

```
p4 <- ggplot(EPL, aes(x = Year, y = PE, fill = Year)) +
  geom_boxplot() +
  theme_minimal() +
  ggtitle("Box Plot of Passing Expected (%) by Year") +
  xlab("Year") + ylab("PE (%)")
p1
p2
p3
p4

# 3. Correlation matrix for WinPCT, DB, PS, PM, PE
cor_matrix <- cor(EPL[, vars], method = "pearson", use = "complete.obs")
print("Pearson Correlation Matrix:")
print(round(cor_matrix, 3))

# 4. Multiple regression of DB, PS, PM, PE on WinPCT
mlr1<-lm(EPL$WinPCT ~ EPL$DB + EPL$PS + EPL$PM + EPL$PE)
summary(mlr1)
print("Multiple Regression Summary:")
```