Applied Mathematics and Nonlinear Sciences

<https://www.sciendo.com>

Research and application of constructing football training linear programming based on multiple linear regression equation

YinZhuang Bai¹, Aizhen Ren^{2†}, Adil Omar Khadidos³, Moaiad Ahmad Khder⁴¹ College of Physical Education, Inner Mongolia Normal University, Hohhot 010020, China² College of Science, Inner Mongolia Agricultural University, Hohhot 010018, China³ Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia⁴ Applied Science University, Al Eker, Kingdom of Bahrain

Submission Info

Communicated by Juan Luis García Guirao

Received June 17th 2021

Accepted September 24th 2021

Available online December 13th 2021

Abstract

Based on the factors affecting sports performance, from a more comprehensive and broad perspective, after consulting the literature, 52 factors that affect the outcome of football matches are selected, including technology, tactics, physical fitness and referees' penalties. By watching the video of the game, 52 influencing factors of 200 games and 400 teams were counted. The original data was statistically processed with correlation analysis and multiple linear regression analysis, and the statistics of the 26 European Cup games were substituted into the winning formula. To verify the scientific nature and objectivity of the formula, we aim to ascertain the core factors in the winning factors of a football game and the quantitative relationship between these factors and the result of the game, so as to provide a certain reference for football training, game analysis and scientific research. The technical and tactical ability of individuals and teams is the core competitive ability factor that affects the result of the game; from a single factor, 15 factor indicators have a significant impact on the result of a football match; on the whole, 10 factor indicators have a significant effect on the result of a football match. In addition, there is a certain quantitative relationship between these influencing factors and the results of the game; empirical evidence shows that the football game winning formula has a certain degree of science and objectivity.

Keywords: football match, winning factors, winning formula, multiple linear regression equation**AMS 2010 codes:** 90C08[†]Corresponding author.Email address: ren_az88@126.com

1 Introduction

Training theory research and football competition practice show that factors such as the competitive ability of their own players and the team's group competitive ability, the opponent's competitive ability and the team's group competitive ability, and the referee's law enforcement ability and moral quality determine the football competition results. The contradictory analysis method provided by the law of the unity of opposites of materialist dialectics enlightens us: Among the many complex factors that determine the result of a football match, some factors have a decisive influence on the result of a football match. These factors are a football team's training and guarantee work. The areas that need to be addressed are identified and improved. Therefore, this study collected 52 indicators reflecting the football team's technical ability, tactical ability, physical fitness and referee's enforcement by watching videos. The original data was statistically processed using correlation analysis and multiple linear regression analysis, and the European Cup was used as the source. The statistical data of 26 games is substituted into the winning formula to verify the scientific nature and objectivity of the formula. It aims to find out the decisive factors that affect the results of football matches and the inherent quantitative relationships between these factors, and is the focus of training for all types of football teams. The determination of football as the target of our analysis and the analysis of player and team performance provide certain theoretical guidance, and also provide new ideas for the research of football researchers [1].

2 Research objects and methods

2.1 Research object

The research object is the core winning factors and winning formulas of football matches.

2.2 Research methods

2.2.1 Document Method

This study mainly used 'football training', 'football match', 'World Cup' and 'European Cup' as keywords to retrieve relevant documents in the CNKI full-text database. After screening, we read 50 new documents in the past 5 years, and at the same time consulted 8 The Ministry's monographs on football training provide documentary support for the determination of factor indicators and statistical scales and the analysis and discussion of statistical results.

2.2.2 Observation method

Watched 200 games of 4 levels of games, and counted 52 factor indicators for 400 teams. These 52 factor indicators reflect the technical ability, tactical ability, physical ability and referee's enforcement in football matches. Among the five elements that reflect the players' competitive ability and the team's competitive ability, psychological and sports intelligence indicators are difficult to measure and are reflected in the technical ability, tactical ability and physical performance of the players and the team, so they are not included in the statistics. The indicators used in the analysis are as follows: The statistics of the game are divided into 4 levels: the first level is the world competition, that is, the 2014 World Cup final stage 16 knockout games; the second level is the intercontinental competition, that is, the 2016 European Cup final stage 15 knockout games and the Americas. There are 8 knockout matches in the final stage of the Cup; the third level is the intercontinental club league, that is, the 2015–2016 season UEFA Champions League 29 knockout matches and the 2016 season Asian Champions League 29 knockout matches; the fourth level is the national league, including 2015 – the top five European leagues in 2016 (15 games each, 75 games in total) and the 2016 Chinese Super League (28 games). Some indicators are calculated using team twelve's football technical statistical software, and some indicators are calculated by watching videos according to statistical standards by dedicated personnel. All competitions only count the number of indicators in the regular game time, and the extra time part is not counted [2].

2.2.3 Statistical method

The SPSS20.0 software package was used to carry out correlation analysis and multiple linear regression analysis on the obtained data. Especially on the basis of the correlation analysis between each factor index and the competition result, the quantitative relationship between each factor index and the competition result is explored by using multiple linear regression analysis. And the statistical data of 26 European Cup matches were substituted into the winning formula to verify the scientific nature and objectivity of the formula.

2.2.4 Logic method

On the basis of using statistical methods to process the original data, the logical methods of analysis, comparison, induction, deduction and reasoning are used to analyse the factors that play a decisive role in the results of football matches. At the same time, the relationship between factor indicators and football results is discussed, which is a quantitative relationship.

3 Statistical indicators and scales

3.1 Statistical indicators

3.1.1 Offensive indicators

These indicators include team possession rate (X1), number of goals (X2), number of shots (X3), number of shots (X4), number of free kicks (X5), number of free kicks (X6), number of free kicks (X7), the number of corner kicks (X8), the number of crosses (X8), the success rate of crosses (X9), the successful entry into the 30 m area of the front field (X10), the successful entry into the penalty area (X11), the number of assists (X12), the number of counterattacks (X13), the number of passes (X14), pass success rate (X15), total number of short passes (X16), total number of mid-range passes (X17), total number of long passes (X18), total number of forward passes (X19), the total number of cross passes (X20), the total number of return passes (X21), the total number of breakthroughs (X22) and the number of offsides (X23).

3.1.2 Defensive indicators

These indicators include the number of goals conceded (X25), the number of shots (X26), the number of shots (X27), the number of steals (X28), the success rate of steals (X29), the number of steals (X30), the number of fights (X31), the success rate of the top (X32), the number of clearances (X33), the number of sieges (X34), the number of blocked shots (X35), the number of blocked shots (X36), the number of blocked passes (X37), the number of yellow cards (X38), number of red cards (X39), number of saves (X40), number of fouls (X41), number of tackles (X42) and success rate of tackles (X43).

3.1.3 Physical fitness indicators

These indicators include full-court running distance (X44), ball control running distance (X45), non-ball running distance (X46), low-intensity running distance (X47), medium-intensity running distance (X48), high intensity running distance (X49) and extreme intensity running distance (X50).

3.1.4 Referee enforcement indicators

These include the number of favourable penalties (X51) and the number of unfavourable penalties (X52).

3.2 Construction of multiple linear regression equation

The independent variable $X_1 - X_{27}$ in the regression model corresponds to each secondary index, and the detailed definition is shown in Table 1. The dependent variable Y corresponds to the total score of the school's education informatisation level. The preliminary analysis of the data shows that the dependent variable and the

Table 1 Stepwise regression process

Step	Introduced variables	Number of variables	Coefficient of determination R ²	Model R ²	C (P)	F	Significance (Pr > F)
1	X6	1	0.7855	0.7855	354.351	183.10	<0.0001
2	X13	2	0.0971	0.8826	174.172	40.54	<0.0001
3	X2	3	0.0342	0.9168	111.994	19.75	<0.0001
4	X20	4	0.0269	0.9437	63.5307	22.47	<0.0001
5	X11	5	0.0174	0.9611	32.8756	20.61	<0.0001
6	X21	6	0.0072	0.9684	21.3579	10.25	0.0025
7	X23	7	0.0044	0.9728	15.0255	7.19	0.0103
8	X19	8	0.0043	0.9771	9.0000	8.03	0.0070

independent variable are roughly linear, so the following multiple linear regression model is established:

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2 + \cdots + b_kX_k \quad (1)$$

where $\hat{Y}$ is the estimated value of the mean of the dependent variable under the condition of the respective variable taking a certain value, $X_1, X_2, \dots, X_k$ are the independent variables, k is the number of independent variables, b_0 is the constant term of the regression equation, also called the intercept, $b_1, b_2, \dots, b_k$ are called partial regression coefficients and b_j represents the average change of Y after X_j changes by one unit under the fixed condition of independent variables other than X_j .

In multiple linear regression analysis, it is necessary to study how many independent variables should be introduced into the model. If fewer independent variables are introduced, the regression equation will not be able to explain the changes of dependent variables well; but it is not that the more independent variables are the better, so it is necessary to adopt some strategies to control and filter the independent variables into the regression equation. We adopt the stepwise regression method, which is to test the import threshold of the P value according to the set regression coefficient significance, introduce independent variables into the model one by one, recalculate the P values of all coefficients in the model and screen variables according to the set elimination threshold. When selecting independent variables, we first select the variable with the highest linear correlation coefficient with the dependent variable into the equation, and perform various tests of the regression equation; then, we find the variable with the highest partial correlation coefficient with the dependent variable and pass the test among the remaining variables [3]. Regression equation, and various tests, are performed on the newly established regression equation; this process is repeated until there are no more variables that can enter the equation. Data processing is performed in SAS 8.01 for Windows software, the significance level of entering the model is set to 0.05 and the significance level of the excluded or retained variables is also set to 0.05. The analysis process of selecting variables and performing regression is shown in Table 1, the statistical results of model regression are shown in Table 2 and the regression coefficients are shown in Table 3.

The regression equation is:

$$Y = 0.074 + 0.131X_2 + 0.277X_6 + 0.129X_{11} + 0.092X_{13} + 0.069X_{19} + 0.074X_{20} + 0.047X_{21} + 0.063X_{23} \quad (2)$$

Table 2 Regression statistics results

	DF	SS	MS	F	Significance (Pr > F)
Regression analysis	8	0.14009	0.01751	229.10	<0.0001
Residual	43	0.00329	0.00007644		
Total	51	0.14338			

DF, degree of freedom; MS, mean square; SS, total square.

Table 3 Regression coefficient table

Variable	Parameter estimate	Standard error	Type II SS	F value	Significance (Pr > F)
Intercept	0.07407	0.01719	0.00142	18.57	<0.0001
X6	0.27741	0.02728	0.00790	103.40	<0.0001
X13	0.09219	0.01663	0.00235	30.72	<0.0001
X2	0.13100	0.02305	0.00247	32.30	<0.0001
X20	0.07443	0.01342	0.00235	30.74	<.0001
X11	0.12904	0.02254	0.00250	32.77	<0.0001
X21	0.04669	0.01905	0.00046	6.01	0.0184
X23	0.06296	0.02080	0.00070	9.17	0.0042
X19	0.06853	0.02419	0.00061	8.03	0.0070

SS, total square.

3.3 Model test

3.3.1 The goodness of fit test of the regression equation (R2 test)

The coefficient of determination $R^2 = 0.9771$, which is very close to 1, indicating that the regression equation is highly significant. $X_2, X_6, X_{11}, X_{13}, X_{19}, X_{20}, X_{21}, X_{23}$ has a highly significant linear effect on Y as a whole (Table 1).

3.3.2 Linearity test of regression equation (F test)

The F test value involves ascertaining whether the independent variable has a significant effect on the dependent variable as a whole. The F test value is 229.10, and the significance (Pr > F) is <0.0001, indicating that $X_2, X_6, X_{11}, X_{13}, X_{19}, X_{20}, X_{21}, X_{23}$ has a significant influence on Y , and the regression effect is very significant (Table 2).

3.3.3 Significance test of regression parameters (t test)

Table 4 shows the estimated value, standard error and other data of the regression coefficient, indicating $X_2, X_6, X_{11}, X_{13}, X_{19}, X_{20}, X_{21}, X_{23}$. All eight independent variables passed the regression coefficient significance test.

3.3.4 Residual error analysis

Residual error refers to the difference between the actual observation value and the regression estimated value. Residual error analysis refers to analysing the reliability, periodicity or other interference of the data through the information provided by the residual error. The calculation formulas for residual and relative error are the following:

$$E = Y - \hat{Y}, \quad \Delta E = \frac{Y - \hat{Y}}{Y}. \quad (3)$$

where Y is the total score of the information level given in Table 1, and $\hat{Y}$ is the predicted value obtained by Eq. (2). The maximum relative error $\max \Delta E < 2.21\%$ is obtained by calculation, which shows that the regression Eq. (2) has high accuracy.

3.4 Statistical scale

Front court 30 m area and side road area, FIFA's standard for field division has been adopted (Figure 1).

For running intensity, we are adopting the China Super League running intensity level standard, that is, extreme intensity running for speeds >21 km per hour; high intensity running for speeds of 17–21 km per hour; medium intensity running for speeds of 14–17 km per hour; and 11–14 km per hour run for low intensity. Other indicators refer to the 2016/2017 FIFA competition rules.

Fig. 1 Division of football match fields.

4 Results and analysis

4.1 Analysis of core winning factors in football matches

Correlation analysis is used to obtain the correlation coefficients between 52 factor indicators and the results of the game (Table 4): total possession rate, number of goals, number of shots, number of shots, number of free kicks, number of successful attempts into the penalty area and number of counterattacks; forward 10 factors that reflect offensive skills and tactics, including: the number of straight passes, the number of cross passes and the number of breakthroughs; as well as 5 factors that reflect defensive skills and tactics, including: the number of goals conceded, the number of shots, the number of shots and the number of sieges and saves. The indicators have a significant relationship with the result of the game. This shows that these 15 technical and tactical factors have a significant impact on the results of football matches, and it also supports the theory that the level of skills and tactics is the core winning factor in determining the results of football matches.

Owning the ball is one of the fundamental ways to score goals in a football game. Therefore, long-term possession of the ball indicates an expansion of the chance of scoring goals. In addition, long-term possession of the ball can not only reduce the player's physical consumption but also increase the opponent's psychological pressure, which creates good conditions for winning the final game. The number of goals scored, the number of shots and the number of shots on target are all indicators that reflect a team's shooting ability and shooting efficiency. The only way to win a football game is to score more goals than the opponent. Therefore, if the team wants to win the game, it must improve the accuracy of the shot while strengthening the shot. Only in this way can the chance of scoring be increased and the game won. Therefore, these three indicators are the fundamental factors that determine the outcome of the game [4].

Table 4 The correlation between the outcome of the game and the indicators

Index	Sample size	Correlation coefficient	Significance	Index	Sample size	Correlation coefficient	Significance
X1	400	0.475	0.047	X27	400	-0.683	0.002
X2	400	0.976	0	X28	400	-0.183	0.469
X3	400	0.535	0.022	X29	400	0.304	0.22
X4	400	0.683	0.002	X30	400	0.122	0.63
X5	400	0.545	0.019	X31	400	0.011	0.913
X6	400	0.416	0.086	X32	400	0.34	0.167
X7	400	0.131	0.604	X33	400	0.452	0.06
X8	400	0.098	0.7	X34	400	0.657	0.003
X9	400	0.024	0.924	X35	400	0.354	0.15
X10	400	0.085	0.737	X36	400	-0.219	0.383
X11	400	0.438	0.069	X37	400	-0.280	0.261
X12	400	0.523	0.026	X38	400	-0.173	0.493
X13	400	0.401	0.099	X39	400	0.427	0.077
X14	400	0.617	0.006	X40	400	-0.527	0.025
X15	400	0.377	0.123	X41	400	-0.401	0.099
X16	400	0.295	0.235	X42	400	-0.450	0.061
X17	400	0.413	0.088	X43	400	0.225	0.369
X18	400	0.402	0.098	X44	400	0.228	0.363
X19	400	0.061	0.809	X45	400	-0.085	0.737
X20	400	0.511	0.03	X46	400	0.182	0.469
X21	400	0.547	0.019	X47	400	0.267	0.284
X22	400	0.206	0.411	X48	400	0.383	0.117
X23	400	0.802	0	X49	400	0.389	0.111
X24	400	-0.012	0.962	X50	400	0.073	0.774
X25	400	-0.976	0	X51	400	0.152	0.548
X26	400	-0.559	0.016	X52	400	-0.152	0.548

Under the guidance of the concept of ‘defines first, win first’, the coach puts the improvement of players and the team’s defensive ability at the top of training. Through daily targeted training, the players’ personal defensive ability and the team’s overall defensive ability have undergone significant improvement, resulting in the time and space available for the use of offensive players to become narrower and more difficult. In this case, as a method of set-kick offence, free-kick offence can use the advantageous conditions of set-kick to complete the shot and get the chance to score. Therefore, the implication is that free kick is also important for the team to win the game in modern football matches.

According to the sports classification method, football belongs to the hit category. Only by approaching the opponent’s goal in the game can the goal rate be increased. Therefore, the restricted area has become key for both the offensive players and the defenders to compete in. Successfully scoring the opponent’s penalty area means the threat of shooting, shooting accuracy and the improvement of the scoring rate, which largely determines the probability of the team winning. Therefore, the number of instances of successfully scoring the opponent’s penalty area has become one of the important indicators that affect the outcome of football matches.

As the coaches of each team attach great importance to defensive capabilities, the individual defensive capabilities of the players of each team have been continuously improved, the team’s partial and overall defensive organisation has become more rigorous, and coupled with the qualitative improvement of the players’ physical level, the offensive team is fighting in position. The time and space available in China is narrower and more difficult. In this case, taking advantage of the defensive team’s loose defensive formation and the unformed

defensive organisation, at the moment of the offensive and defensive transition, to launch a quick counterattack has become one of the important methods of each team's offensive, and it has also become the main attack method for scoring goals. Therefore, counterattack tactics are the main offensive means for teams to win games in modern football matches.

One of the outstanding features of modern football games is the high level of integrity. Therefore, the continuous improvement of the overall defensive ability of each team has made it more and more difficult for individual offences. Relying on the overall offence of the team is an important means to win the game. Therefore, passing has become the main means of connecting players and completing team offences. In the passing direction, a forward pass can break through the opponent's defensive line and create a chance to score a goal; while a cross pass can shift the offensive direction, mobilise the opponent's defensive line, create vertical penetration space and form a local number advantage. Therefore, these two types of passes greatly affect the outcome of the game.

Under the situation that each team attaches great importance to defence, the players' personal defensive ability and the team's local and overall defensive organisation ability have been greatly improved. In this case, only the player's personal breakthrough ability can be continuously improved. Only in the confrontation with the defensive players can the victory be won, by breaking through the opponent's defence, thereby forming an advantage in the number of offensives in a local area, creating favourable conditions for finally breaking through the opponent's defensive line, facilitating availability of optimal shooting opportunities and completing the goal. Therefore, personal breakthrough ability also has an important impact on the outcome of the game [5].

The number of goals conceded, the number of shots and the number of shots taken together reflect the final effect of the team's defence. In the course of the game, the team can only rely on the excellent personal defensive ability of the players and the local and overall tight defensive formation to make it difficult for the offensive team to complete the shot. In particular, to ensure that the goal is not lost, the team can win the game on this basis. Therefore, the number of goals conceded, the number of shots and the number of shots are important indicators that affect the outcome of the game.

Encirclement is the main indicator reflecting the local defensive organisation. In the game, through the first defender's frontal delay or close pressing against the team player, combined with the partial coordination, orderly cooperation and rapid movement of the second and third defenders, they form an organisation with the first defender. Strict and well-defined local defensive formations form an encirclement to the team members, thereby improving the team's defensive quality and ensuring that the team concedes fewer or no goals. Therefore, the level of siege directly affects the quality of the team's defence, and to a large extent determines the outcome of the game.

As the last defender of the team, 'a good goalkeeper is equal to half a team' has become a consensus in football. A save is an important indicator that reflects the goalkeeper's defensive ability. The strength of this ability to save plays a very important role in ensuring that the team does not concede a goal. Therefore, improving the goalkeeper's ability to save has an important impact on the result of the game.

The high correlation between the above 15 offensive and defensive technical and tactical factors and the results of the game enlighten us to the following inference: deep understanding of modern football concepts, following modern football development trends, strengthening the players' ball control and breakthrough capabilities, improving the quality of the team's direct and cross passes, making full use of the advantages of counterattacks and free kicks to launch an offence and a good grasp of shooting opportunities is where coaches need to focus on improving offensive training in the future. At the same time, it is necessary to improve the team's partial and overall defensive organisation ability, reduce the offensive team's shots and try to avoid the team's conceding. In addition, the importance of goalkeepers has become increasingly prominent, and goalkeeper training should be strengthened [6].

4.2 Discussion on the winning formula model of football matches

Through correlation analysis, we only understand the correlation between each factor index and the game result, but it cannot reflect the relationship between each factor index and the game result as a whole. In order to grasp and understand the impact of various factors on the results of the game as a whole, and to explore the quantitative relationship between the two, a multiple linear regression analysis was carried out with the game results as the dependent variable and each factor index as the independent variable.

The regression analysis adopts the full entry method. The correlation coefficient value (R) is 0.882, the determination coefficient R^2 (Square) is 0.778 and the adjusted R^2 (Adjusted Square) is 0.578, indicating that the regression model has a high degree of fit, and 10 selected independent variables: the number of goals, the number of free kicks, the number of successful attempts into the penalty area, the number of counterattacks, the number of forward passes, the number of cross passes, the number of breakthroughs, the number of goals conceded, the number of sieges and the number of saves explain 57.8 of the variance in the total score percentage, and the Durbin–Watson test value is 2.270, indicating that the residuals are independent of each other. The Durbin–Watson test is passed, and the regression equation can be used. The results of the analysis of variance calculated by the multiple linear regression analysis show (Table 5) that, in this multiple linear regression model, F is 3.888, $P < 0.01$ and there is a very significant difference, indicating that the regression analysis is meaningful [7].

Table 5 Regression analysis variance of game winning factors and game performance

Model	Sum of variance	Mean difference	F	Sig.
1	Regression	129.495	2.878	3.8880.000
	Residual	37.005	0.74	
	Total	160.5		

Given the significance level $\alpha = 0.05$, the P value of a goal is 0.029, the P value of a free kick is 0.016, the P value of a successful penalty area is 0.021, the P value of a counterattack is 0.008 and the pass is forward. The P value for cross passes is 0.003, the P value for cross passes is 0.038, the P value for breakthroughs is 0.033, the P value for conceded goals is 0.000, the P value for surrounds is 0.045 and the P value for saves is 0.048. It shows that the 10 variables introduced by the regression model have a significant impact on the results of the game. They passed the significance test, and the constant P value was 0.654, which failed the significance test. The following can be observed from the results of the regression equation: the result of the football match and the number of goals scored by the team, the number of free kicks, the number of successful attempts into the opponent's penalty area, the number of counterattacks, the number of forward passes, the number of cross passes, the number of breakthroughs and the number of sieges. There is a positive correlation among the eight technical and tactical indicators, namely, the number of goals scored, the successful entry into the penalty area, the free kick, the counterattack, the forward pass, the cross pass, the breakthrough and the siege; the higher the value of these indicators, the higher the probability of the team winning. The result of the game is negatively correlated with the number of goals conceded and the number of saves, that is, the fewer goals conceded and the fewer saves, the better the team's performance may be.

What needs further explanation is that of the 15 factor indicators that are highly correlated with the game results through correlation analysis, only 10 indicators in the multiple linear regression model have high correlation with the game results. The reason is the unselected shots. There is a high correlation between the number of shots and the number of goals scored, so the most representative number of goals is selected; the number of goals shot and the number of shots are highly correlated with the number of goals conceded, so the most representative conceded goals were incorporated in the number selected; and the ball possession rate was not selected, mainly because the ball control is mainly reflected by the passing of the team's offence and the breakthrough of the individual offence, so there is a high correlation with the passing and the breakthrough, forward passes and cross passes. The selection of the ball and breakthrough largely reflects and represents the ball possession rate.

In order to verify whether the winning formula has a certain degree of scientific and objectivity, we take the 26 games in which the win–loss relationship is determined within 90 min of the group stage of the 2016 European Cup finals as an example, the 26 games and 52 team goals etc. The core winning factor indicators are substituted into the winning formula. Among the 22 games, the team with the larger Y value calculated by the winning formula won, and the winning rate reached 85%, which shows that the winning formula has a certain degree of scientific temperament and objectivity [8].

In terms of specific competition cases, the two matches of Wales vs. Slovakia (the result of the two sides' match is 2:1) and England vs. Wales (the result of the two sides' match is 2:1) are randomly selected as examples. The relative ratios of the core winning factors of Wales and Slovakia, England and Wales are replaced by the equation. The calculation result is: the Y value of the Wales team is 0.22638, the Y value of the Slovak team is 0.15862, the sum of the two Y values is 0.38500, and the calculated Welsh relative winning percentage of the team is 58.80%, while the relative winning percentage of Slovakia is 41.20%; the Y value of England is 0.32222, the Y value of Wales is 0.06218 and the sum of the two Y values is 0.38440 [9, 10]. The calculated relative winning percentage of England is 83.82%, and the relative winning percentage of the Wales team was 16.18%. From the calculation results of the winning formula, it can be observed that Wales and England have a higher winning percentage than their respective opponents Slovakia and Wales, and the results of the game also prove that the team with a relatively high winning percentage has won the game, which further proves that the winning formula has a certain degree. The scientific and objectivity of It is important to note that the regression formula model obtained by this research only reflects the objective quantitative law between the football game and the core winning factors to a certain extent. Strictly speaking, the football game result is affected by multiple subjective and objective factors. There is no absolute quantitative relationship between the result of the game and the factors. This formula model is only a discussion. In the future, more in-depth discussions are required from football researchers.

5 Conclusion

The technical and tactical capabilities of individuals and teams are the core competitive ability factors that affect the outcome of the game. The number of goals scored, the number of free kicks, the number of successful attempts into the penalty area, the number of counterattacks, the number of forward passes, the number of cross passes and the number of breakthroughs are indicators of offensive factors; the number of goals conceded, the number of sieges and saves are indicators of defensive factors that affect the core winning factor of football match results. There is a certain quantitative relationship between the 10 core winning factors and the results of football matches. The quantitative relationship in this research is only intended to provide a formula model to reflect the quantitative law between them. Empirical evidence shows that the winning formula of football matches is scientific and objective.

References

- [1] Dijana P D. Measuring efficiency of football teams by multi-stage DEA model[J]. *Tehnicki Vjesnik*, 2015, 22(3):763-770.
- [2] Viano D C, Withnall C, Halstead D. Correction to: Impact Performance of Modern Football Helmets[J]. *Annals of Biomedical Engineering*, 2018, 46(1):208-210.
- [3] Wilmes E, Ruiter C, Bastiaansen B, et al. Inertial Sensor-Based Motion Tracking in Football with Movement Intensity Quantification[J]. *Sensors*, 2020, 20(9):2527.
- [4] Pezdek K, Michaluk T. The functioning of the Polish Football Association from the perspective of Michel Foucault's conception of exclusion[J]. *Journal of the American Chemical Society*, 2016, 81(4):1-14.
- [5] Constantinou A, Fenton N. Towards Smart-Data: Improving predictive accuracy in long-term football team performance[J]. *Knowledge-Based Systems*, 2017, 124(MAY15):93-104.
- [6] Siegmund G P, Guskiewicz K M, Marshall S W, et al. Laboratory Validation of Two Wearable Sensor Systems for

- Measuring Head Impact Severity in Football Players[J]. *Annals of Biomedical Engineering*, 2016, 44(4):1257-1274.
- [7] Campbell K R, Warnica M J, Levine I C, et al. Laboratory Evaluation of the gForce Tracker, a Head Impact Kinematic Measuring Device for Use in Football Helmets[J]. *Annals of Biomedical Engineering*, 2016, 44(4):1246-1256.
- [8] Matias A, Jonthan T G. Computer Football: Plays, Players and Strategies Choices[J]. *IEEE Latin America Transactions*, 2018, 16(5):1485-1492.
- [9] Dusunceli F. New Exact Solutions for Generalized (3+1) Shallow Water-Like (SWL) Equation[J]. *Applied Mathematics and Nonlinear Sciences*, 2019, 4(2):365-370.
- [10] Ahin R, Yagci O. Fractional Calculus of the Extended Hypergeometric Function[J]. *Applied Mathematics and Nonlinear Sciences*, 2020, 5(1):369-384.

This page is intentionally left blank