Original Article

Using Poisson model for goal prediction in European football

TUGBAY INAN

Faculty of Sport Sciences, Dokuz Eylul University, Izmır, Turkey

ABSTRACT

Predicting the features of behaviour of big data and multivariable systems has been a research subject in various fields of science. When it comes to football, as it is a field of sports followed by the whole world, the number of studies carried out aiming at predicting the results of football games has been increasing in the field of football science. Although the result of a football match depends on various variables, it is mainly determined over the offensive and defensive strengths of the teams. Different variables have so far been determined in the literature to figure out these strengths of the teams. In this study, it was aimed to predict correctly how many goals a team could score or concede in the last 5 weeks based on the average number of goals they scored and conceded since the beginning of the season in 6 European leagues. For this reason, a Poisson distribution model was established based on these offensive and defensive strengths. A total of 4264 matches and 5938 goals were analysed in the study and the established model yielded affirmative results at the level of 50% in the leagues analysed.
Keywords: Poisson Model; Goal; Football; Soccer; Prediction; European leagues.

Cite this article as:
Inan, T. (2021). Using Poisson model for goal prediction in European football. *Journal of Human Sport and Exercise*, *16*(4), 942-955. https://doi.org/10.14198/jhse.2021.164.16

Corresponding author. *Department of Physical Education and Sports, Faculty of Sport Sciences, Dokuz Eylul University, Izmır, Turkey. https://orcid.org/0000-0002-3573-8086*
E-mail: tugbay.inan@deu.edu.tr
Submitted for publication February 11, 2020.
Accepted for publication April 11, 2020.
Published October 01, 2021 *(in press* June 19, 2020)
JOURNAL OF HUMAN SPORT & EXERCISE ISSN 1988-5202

doi:10.14198/jhse.2021.164.16

INTRODUCTION

Today, football, which we increasingly encounter in larger and more advanced organizations as a sports event, has also evolved into a crowd puller around the world. In recent years, with the latest technological developments, football has its name and consent frequently mentioned as a field of sports about which all information has become accessible. These developments have initiated a structural transformation process in the football sector all over the world. With this change over the time, football has gradually evolved into an industry needing to be managed based on the principles of maximization of profit and benefit through strategic and financial techniques. Today, football organizations around the world have taken the form of a profitable business. The reports published following the studies conducted on football economy estimate that despite the effects of the global economic crisis, the European football market is still growing economically and by the end of the 2017-2018 season, this size is estimated to reach 27 billion Euros. According to the report published by Deloitte, the income of 20 clubs with the highest income in Europe was determined as EUR 9.3 billion in 2017.

This economic size and development in the football field have influenced the technical-tactical understanding of the football game, the players' characteristics, the types of fans, the social and economic levels of those interested in football, which has led to structural developments in the management perception of football clubs. This change and development in football have attracted the attention of different disciplines and researches have been carried out in an academic sense on the field. Some of these researches are composed of econometric researches. With the help of mathematical models and the science of statistics, different studies are frequently encountered in the literature on this field.

In recent years, the models which attempt to explain the variables which will have an effect on football match results have been started to be used frequently. We can say that the most important factor related to the reason why football is so popular around the world is the fact that the results of the matches cannot be predicted beforehand.

In particular, if all conditions are equal for the home and visiting team, the fact that the match result is not known beforehand creates a great interest for all people interested in football and also increases the desire to follow the matches. This problem has led to the development of methodologies aimed to predict the match results using scientific methods before the games are played. When the literature is examined, it can be seen that the factors affecting the match result have already been determined by the majority of the studies and that mathematical models and statistical methods which calculate the impact shares of these factors and then classify them according to their functions have been used. The majority of the statistical methods consist of models that calculate what is expected in the next game by using the regression technique over the averages of the variables that can have an effect on the match result from the beginning of the season or in the last week within a given past time period. In fact, there are many variables that affect the results of football matches. Some of these are related to the internal dynamics of the tools and are measurable. However, for the visiting team, the variables that affect the result of the match in subjective examples are such ones as weather condition, location, team morale, coach tactics. Therefore, the result of the match depends on many factors, so that it is necessary to reduce these factors to measurable numerical values. Variables such as points obtained in and out, number of goals scored, number of dangerous shots thrown from the penalty area should be made meaningful via mathematical models and the offensive and defensive strengths of the teams should be calculated.

When the literature is examined, we encounter different methods to calculate the strengths of the teams based on various variables. On the other hand, it can be seen that most of the studies were carried out by linear regression analysis. For the first time in the literature, Moroney (1956) pointed out that the predictions of football matches could be explained with a statistical model. The book states that the negative binomial distribution and the Poisson distribution are appropriate methods to be utilized in predicting the results of football matches.

As with all professional team sports, it is possible to say that there is an advantage in football for the home teams (Pollard 1986). There are many studies in the literature pointing out that it is advantageous to play at home field. According to Pollard (1986), in the UK League, the home field advantage decreased from 67.9% to 63.9%. Courneya and Carron (1992), in their study they carried out for home teams, found the percentage of home teams winning the game as 69%. Following this study, Carron and Hausenblas (1998) reported that the home field advantage is in favour of the home team. Pollard and Pollard (2005) stated that home field advantage values in most countries where European football is played are between 60% and 65%. They also found that there are multiple variables that affect the home field advantage as a result of the studies they conducted in line with home field advantage hypothesis.

It has been revealed in almost all of the studies based on the home field advantage that the result of a football game is determined by mathematical calculations according to the league table formed at the end of the season. The basis of these studies are formed by the calculation method proposed by Pollard (1986). However, this method shows us only whether there is home field advantage, and if so, what the percentage of this advantage is in terms of the team. In time, especially in football, as a result of the increase in the interest in predicting the results of the matches, different studies in the field of sports science have shown up. It is especially vital to be able to predict the results of matches both for betting companies and bettors. Two systems are used aimed towards betting games in football. The first one is referred to as paramutual games. According to this system, the bettors will not know how much they will win, even if they have correctly predicted the results of matches. This is because all the money deposited by the bettors is collected in a specific pool and is paid to the winners after the deductions have been made from the people who predicted the results of the matches. There are a number of risks for the better.

For weekly match results, money is streamed to the system by people who bet on certain amounts and the money collected is shared by the people who have predicted the results correctly. If the predictions are not made correctly, all the money bet is left to the company. There are no commercial risks posed for the betting company in this system. As a result of this, many people are needed to stream money to the system by depositing large amounts of money for those betting in this system to make a great deal of money, and also very few people should predict the results of the matches correctly.

While in the other system, there are some risks for the betting company. In the system called Fix odd betting, the result of the match must be explained by the company beforehand. In this system, the bettor will know in advance how much money he will earn if he correctly predicts the result of the match before the matches are completed. There is no risk of the person placing a bet here; however, the betting company must correctly predict the result of the matches, otherwise the odds to be misrepresented will increase the commercial risk, even leading the company to pay more than the money it has collected.

It is observed that two-variable Poisson and probity regression models were used in the studies to predict the results of the matches. Crowder et al. (2002) analysed the British Premier League matches between 1992 and 1997 and discussed the dynamic models between a team's scoring and conceding a goal. Moura et al.

(2007) analysed Brazilian First League goals based on the possible goal shots. Leitner et al. (2010) carried out studies aiming to determine the degree of impact of various variables such as corner shots, free kicks and offside positions. Williams and Walters (2011) examined the effects of altitude above sea level in football matches in South America. Koopman and Lit (2012) expressed in their study where they benefited from a two-variable dynamic Poisson model that the number of goals scored by a team indicates the attacking power of the team and that it depends on the defence capacity of the opponent, home field advantage, if any, and the importance of the match. They tried to predict the results of the matches according to the Poisson model they formed. In this study, Poisson distributions were calculated to determine the defence and offense based strength of the teams in football and to predict the results of their future matches. For this purpose, this model was established based on the number of goals they scored and conceded by the teams in past matches.

METHODS

Poisson model was used in the study to predict the results of football matches.

Poisson Model

Poisson distribution is a discrete probability distribution in the probability theory and statistics science branches and it aims to express the probability of occurrence of certain events in a fixed time unit interval. It is accepted that the mean number of events occurring during this time period is known and that the time difference between any event and the event that follows it is independent of the previous time differences. While the Poisson distribution is applied to problems with certain fixed time units interval, it can also be successfully applied to football match results. This probability distribution was first introduced by Siméon-Denis Poisson in 1838.

The general focus of the Poisson distribution is a variable event; this event occurs at a certain time interval and the number of events observed in this range is considered to be a random variable for the Poisson distribution. The expected value of the number of events occurring in this fixed range (the mean number of occurrences) is fixed as λ, and this mean value is proportional to the range length. To explain this with the result of the goals scored by the home team inside; If it scored an average of 15 goals in the last 6 matches it has played home, it would have the probability of scoring an average of 20 (= 8x15 / 6) goals in a fixed 8 match intervals. The probability of occurrence of a k-number (k = 0, 1, 2, 3 ...) of any non-negative phenomenon is expressed as follows:

$$f(k,\lambda) = \frac{\lambda^{k} e^{-\lambda}}{k!} \qquad (1)$$

e, the base of the natural logarithm.
k, The number of occurrences of an event whose probability is given with function.
k!, Factorial for k.
λ The expected value of occurrence of an event in a given fixed interval, a positive real number.

The function of k is the probability mass function for the Poisson distribution. The λ parameter for the Poisson distribution is not only the expected value, i.e. the average for events occurring (k) times. At the same time, it can also be probability distribution variance.

$$\sigma^2=\sum[k^2.P(x)] - k^2 \qquad (2)$$

While the mean value of the number of events observed is expressed with, λ the Poisson model having standard deviation and variance.

$$\sigma^2=\sqrt{\lambda} \tag{3}$$

A Poisson distribution can be applied in systems where a large number of events are possible to occur yet where these occurrences are accepted as very rare to occur. In terms of football, what exemplifies this situation best is the example that the possibility of every shot thrown to the goal to be scored is very low. Many conditions must be fulfilled in order to be able to score a goal in a football match and this can be expressed mathematically as a rare number. Then, the number of rare events occurring is expressed as the result of discrete trials. It can also be modelled using a binomial distribution to give more accurate results. But a binomial distribution with n and λ / n parameters (i.e. the probability distribution for (n) number of shot attempts with λ / n success probability for each goal) is more compatible with a Poisson distribution with the expected value of λ as the (n) number of trials approaches to infinite limit. This limit is sometimes referred to as a rare event rule. This expression is somewhat misleading; there can also be some events that can be modelled with many Poisson distributions (such as being more than two goals in one match) which are not at all rare. However, the calculation of the binomial distribution for large numbers requires the use of factorial numbers. This long calculation takes a lot of time and therefore the Poisson distribution in the literature is used instead of approximate binomial distribution.

Subtracting the limit Poisson distribution mass function from binomial distribution is mathematically proved as follows:

The limit used for the variables is calculated as follows.

$$\lim_{n\to\infty}\left(1+\frac{\lambda}{n}\right)^n = e^{-\lambda} \tag{4}$$

p = If λ / n equality is taken into this expression, the following general equation is obtained:

$$\lim_{n\to\infty} Pr(X=k) = \lim_{n\to\infty}\binom{n}{k}p^k(1-p)^{n-k} = \lim_{n\to\infty}\frac{n!}{(n-k)!k!}\left(\frac{\lambda}{n}\right)\mathrm{k}\left(1-\frac{\lambda}{n}\right)^{n-k} \tag{5}$$

When this equation is opened, the result is expressed as follows:

$$= \lim_{n\to\infty}\binom{n}{k}\left(\frac{\mathrm{n-1}}{n}\right)\left(\frac{\mathrm{n-2}}{n}\right)\dots\left(\frac{\mathrm{n-k+1}}{n}\right)\left(\frac{\lambda^k}{k!}\right)\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k} \tag{6}$$

As a result, the following occurs at the limit:

$$\frac{\lambda^{\mathrm{k}}e^{-\lambda}}{k!} \tag{7}$$

More generally, if a row of binomial expressions for binomial random variables with n and pn parameters is as follows,

$$\lim_{n\to\infty} np_{n=\lambda,} \tag{8}$$

This series approximates the series for a Poisson random variable with a mean λ in the distribution.

The expected value and variance value for a random variable exhibiting the Poisson distribution is λ. The high moments of the Poisson distribution are Touchard polynomials formed by the terms λ. If the expected value for the Poisson distribution is 1, then the number of a set of size n equals the n-th moment of that distribution according to Dobinski's formula. The mode value of a random variable that indicates the Poisson distribution with an integer λ lambda parameter equals to the largest positive integer less than λ, namely to λ.

The sum of the random variables showing the Poisson distribution: If the expression of x, ~ Poi (λi) shows Poisson distribution with λi parameter and the terms x, are independent,

$$Y = \sum_{i=1}^{N} X_i \sim Poi\left(\sum_{i=1}^{N} \lambda_i\right) \quad (9)$$

Then, the above given equation shows a Poisson distribution made up of parameter sums whose parameters are added to the sum.

The moment-generating function of the Poisson distribution with the expected value λ is expressed as follows;

$$E\left(e^{tX}\right) = \sum_{k=0}^{\infty} e^{tk} f(k;\lambda) = \sum_{k=0}^{\infty} e^{tk} \frac{\lambda^k e^{-\lambda}}{k!} = e^{\lambda\left(e^{t-1}\right)} \quad (10)$$

For the Poisson distribution, all cumulants are equal to the expected value λ. For Poisson distribution, the N-th factorial moment is λn. Poisson distributions are also infinitely divisible probability distributions and are expressed as follows:

$$X_1 \sim Pois\left(\lambda_1\right) \quad (11)$$

The Poisson distribution can be used to predict the results of sports events, especially for football matches. The variables that are expected to be distributed must be equal to the home and visiting team. Then, the following model emerges:

$$Y = \sum_{i=1}^{N} X_i \,.\, Poi\left(\sum_{i=1}^{N} \lambda_i\right). Poi\left(\sum_{i=1}^{N} \mathrm{k}_i\right) \quad (12)$$

Models predicting the results of a football match should take the differences between the variables of the two teams to compete into account. The excess of variables that can affect the results of football matches cause various problems during the application. For this reason, it is more meaningful to estimate the result of the match based on the goal scored and the goal conceded at home field. The most important feature that distinguishes a strong team from a weak team can be expressed as scoring more goals and conceding less goals.

While the offensive strength of a team suggests how many more and how many less goals a team has scored per game out of the total average in the league they are competing in defensive strength indicates how many more and how many less goals a team has conceded per game out of the total average. When calculating offensive and defensive strengths of teams, the average goal a team has scored per game in the league must be calculated using equation (13).

$$LOd = LATGS/(LTS.N) \tag{13}$$

LOd – The mean variable in the league.
$LATGS$- The total goal scored in the league.
LTS – The number of teams in the league.
N – The number of weeks in evaluation.

Then the offensive and defensive strengths of the teams are calculated. Equation (14)

$$O_i = OAG_i/LOd)$$
$$D_i = OYG_i/LOd) \tag{14}$$

LOd - Number of goals scored/conceded by a team in the league.
OAG_i- i average goal scored by the team.
OYG_i- i the average number of goals the team has conceded.
D_i - i defensive strength of the team.
O_i - i offensive strength of the team.

If the score of a match played between i and j teams is (Xij ,Yij), Xij, which is the number of goals scored by the team (i) against team (j), depends on the offensive strength of i and defensive strength of team j. Similarly, Yij, which is the number of goals scored by team j against team i depends on defensive strength of team i and offensive strength of team j. The function of the teams based on offensive and defensive strengths shows Poisson distribution and the algorithm is given in the equation (15).

$$X_{ij} \sim Poisson\ (LOd.\mathrm{O}_i.\mathrm{D}_j)$$
$$Y_{ij} \sim Poisson\ \left(LOd.\mathrm{D}_i.\mathrm{O}_j\right) \tag{15}$$

LOd – Mean variable in the league.
O_i- i Offensive strength of the team (i).
D_i - i defensive strength of the team (i).
O_j - j offensive strength of team j.
D_j - j defensive strength of team j.

Material

In this study we conducted, a data set made up of the competitions between the last 29th and 38th weeks of the 2017-2018 season of the 6 major leagues in Europe was used (England Premier League, France Ligue 1, Germany Bundesliga, İtaly Serie A, Spain La Liga, Turkey Super League). The dataset was generated from data obtained from the website "*mackolik.com*". In this study, a total of 5938 goals were analysed, with a total of 4264 matches and the total number of goals scored by the teams in all matches. The values of the leagues analysed are shown in Table 1.

Table 1. Summary of leagues.

	Premier League	**Ligue 1**	**Bundesliga**	**Serie A**	**La Liga**	**Super League**
Weeks	34-38	34-38	29-34	34-38	34-38	29-34
Number of matches	760	760	612	760	760	612
Number of goals	1018	1024	855	1017	1024	1000

RESULTS

In the study, a data set consisting of the competitions played in the last 5 weeks of the league was used in 2017-2018 season of the 6 major leagues in Europe. Table 1 shows the week in which the data set was analysed and how many matches were analysed. In this study, Poisson distribution was used to estimate how many goals were scored as a result of a football match. In establishing the model, the scoring strengths of the teams in a league was determined. In the event that the teams involved in the study were home and the away team, the offensive and defensive strengths were given in the appendix. In order to determine the strengths of the teams, the distributions of the goals they conceded and scored during certain weeks in a particular league were used. Poisson model was developed to predict the results of the next weeks' match results through these distributions and strengths. After the establishing Poisson model, the average number of goals and the number of goals we expected to be scored by the home team were determined. The results of the match week were compared with the number of goals in the match in the following week. The percentage of goals predicted correctly according to the Poisson probability results for the last five weeks of the leagues is shown in Table 2.

Table 2. The percentage of goals predicted correctly.

England Premier League	France Ligue 1	Germany Bundesliga	Italy Serie A	Spain La Liga	Turkey Super League
54%	62%	60%	52%	56%	60%

According to the findings obtained from the study; Poisson probability model, which was established in the last five weeks of the leagues, predicted the number of goals to be scored at the highest level in France Ligue at a level of 62%, whereas Italy was the lowest in Serie A with 52%. The established model achieved a success rate of more than 50% in all leagues, as can be seen in Table 2, which shows us that the Poisson probability method is a method that can be used to predict the number of goals to be scored in football matches in the coming weeks. The tables belonging to the last week's results of the findings obtained in the study are given in the appendix.

DISCCUSSION AND CONCLUSION

In 2015, Mwembe et al. conducted a study to determine the results of the matches. They established a model to predict the results of the matches played in the 1st, 16th and 30th weeks of the league and the success rate of the model was determined to be 45.83%.

On the other hand, the fact that the model established in the study consisted alone of the goals scored and conceded correctly predicted the match results in an average of 57.3%. Although this ratio seems to be larger than the examples given above, further studies are needed to increase this rate. Working with a larger data set that will cover the entire league instead of the last week data of the leagues can improve the reliability of the study. In addition, when determining the offensive and defensive strengths of the teams, many factors that affect the results of the match as well as the goals they have scored or conceded must be taken into consideration. Such variables as the passes made, the balls thrown into the penalty area, the shots thrown to goal area should be included into the studies.

In addition, supporting the studies with such algorithms as the home field advantage, the competitive balance index within the league and machine learning would improve the prediction of the number of goals to be scored in the match more accurately. As a result, in addition to determining the number of goals scored in the matches played by the teams in an attempt to achieve higher modelling success, it is thought that the evaluation of the leagues with different variables to be added to the model and the development of a probity regression model by adding different variables will also increase the rate of correct prediction rate of the match results in the future studies.

SUPPORTING AGENCIES

No funding agencies were reported by the author.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the author.

REFERENCES

Carron, A.V., &. Hausenblas, H.A (1998). Group Dynamics in Sport. 3rd Edn., Morgantown, WV: Fitness Information Technology.

Crowder, M., Dixon, M., Ledford, A., & Robinson, M., (2002). Dynamic Modeling and Prediction of English Football League Matches for Betting. The Statistician, 51(2), 157-168. https://doi.org/10.1111/1467-9884.00308

Deloitte, Football Money League 2020. Sports Business Group. https://www2.deloitte.com/uk/en/pages/sports-business-group/articles/deloittefootball-money-league.html (2020, accessed 05 March 2020).

Koopman, S. J., & Lit, R. (2012). A Dynamic Bivariate Poisson Model for Analysing and Forecasting Match Results in the English Premier League. Tinbergen Institute Discussion Paper. Amsterdam. https://doi.org/10.2139/ssrn.2154792

Leitner, C., Zeileis, A.,& Hornik, K. (2010). Forecasting Sports Tournaments by Ratings of (prob) Abilities: A Comparison for the Euro 2008, International Journal of Forecasting, 26(3), 471–481. https://doi.org/10.1016/j.ijforecast.2009.10.001

Moroney, M. J. (1956). Facts From Figures. 3rd edition, Penguin Books , London.

Moura, F. A., Santiago, P. R. P., Misuta, M. S., Barros, R. M. L., & Cunha, S. A. (2007). Analysis of the Shots to Goal Strategies of First Division Brazilian Professional Soccer Teams, in 'ISBS-Conference Proceedings Archive', Vol. 1. Retrieved from https://ojs.ub.uni-konstanz.de/cpa/article/download/483/423/0

Mwembe, D., Sibanda, L., & Mupondo, C. N. (2015). Application of a Bivariate Poisson Model in Devising a Profitable Betting Strategy of the Zimbabwe Premier Soccer League Match Results. American Journal of Theoretical and Applied Statistics. 4(3),99-111. https://doi.org/10.11648/j.ajtas.20150403.15

Pollard, R., (1986). Home Advantage in Aoccer: A Retrospective Analysis. Journal of Sports Sciences, 4(3), 237-248. https://doi.org/10.1080/02640418608732122

Pollard, R., & Pollard, G. (2005). Home Advantage in Soccer: A Review of Its Existence and Causes. International Journal of Soccer and Science Journal, 3(1), 28-38.

Williams, T., &Walters, C. (2011), The Effects of Altitude on Soccer Match Outcomes, In Proceedings of the MIT Sloan Sports Analytics Conference Boston. Retrieved from http://www.sloansportsconference.com/wp-content/uploads/2011/08/The-Effects-of-Altitude-on-Soccer-Match-Outcomes.pdf

Appendix

Table 3. Home, away offensive, and defensive strengths in England.

Team	Home		Away	
	Offensive Strengths	Defensive Strengths	Offensive Strengths	Defensive Strengths
Manchester City	2.119	0.643	2.128	0.475
Manchester United	1.356	0.436	1.375	0.658
Tottenham	1.284	0.582	1.558	0.693
Liverpool	1.504	0.485	1.788	0.970
Chelsea	1.042	0.735	1.548	0.695
Arsenal	1.876	0.919	0.919	1.133
Burnley	0.550	0.727	0.917	0.762
Everton	0.973	1.011	0.726	1.206
Leicester City	0.869	1.011	1.306	1.206
Newcastle United	0.660	0.824	0.825	1.039
Crystal Palace	0.990	1.309	0.733	0.970
Bournemouth	0.903	1.378	0.823	1.097
Watford	0.938	1.424	0.823	1.170
Brighton & Hove Albion	0.834	1.149	0.484	0.914
West Ham United	0.770	1.212	1.100	1.455
Huddersfield Town	0.587	1.164	0.550	1.143
Southampton	0.733	1.212	0.779	1.039
Swansea City	0.587	1.067	0.504	1.108
West Bromwich	0.730	1.332	0.484	0.914
Stoke City	0.695	1.378	0.629	1.353

Table 4. Home and away offensive and defensive strengths in France.

Team	Home		Away	
	Offensive Strengths	Defensive Strengths	Offensive Strengths	Defensive Strengths
Paris Saint-Germain	2.420	0.665	1.780	0.510
Monaco	1.625	0.754	1.639	1.021
Lyon	1.277	0.749	2.174	0.863
Marseille	1.533	0.842	1.597	0.967
Rennes	0.730	0.983	1.287	0.760
Nice	1.003	1.020	1.030	0.948
Bordeaux	0.934	1.020	1.030	0.911
Saint-Etienne	0.985	1.264	0.666	0.794
Montpellier	0.692	0.842	0.703	0.474
Nantes	0.511	0.796	0.932	0.829
Guingamp	1.037	1.064	0.796	1.203
Amiens SC	0.899	0.842	0.468	0.765
Dijon	1.204	1.264	0.887	1.554
Angers	0.795	1.242	0.843	0.802
Lille	0.830	1.197	0.796	1.276
RC Strasbourg	0.934	1.197	0.796	1.421
Caen	0.620	0.889	0.444	1.139
Toulouse	0.766	0.936	0.666	1.139
Troyes	0.620	0.889	0.666	1.278
Metz	0.584	1.545	0.799	1.347

Table 5. Home and away offensive and defensive strengths in Germany.

Team	Home		Away	
	Offensive Strengths	Defensive Strengths	Offensive Strengths	Defensive Strengths
Bayern München	2.166	0.588	1.812	0.482
Schalke 04	1.024	0.802	1.309	0.816
Borussia Dortmund	1.483	1.056	1.230	0.906
Hoffenheim	1.379	0.802	1.410	1.186
Bayer Leverkusen	1.024	0.909	1.460	0.927
RB Leipzig	1.260	1.308	0.909	0.985
Eintracht Frankfurt	0.964	0.956	1.016	0.985
Stuttgart	0.667	0.453	0.749	1.024
Mönchengladbach	1.038	1.006	0.963	1.182
Hertha Berlin	0.827	1.122	1.007	0.704
Augsburg	0.890	1.207	1.016	0.788
Werder Bremen	0.741	0.855	0.802	0.867
Hannover 96	1.038	1.258	0.749	1.024
Mainz 05	0.827	1.016	0.805	1.149
Freiburg	0.591	0.909	0.755	1.446
Wolfsburg	0.748	1.283	0.654	0.853
Hamburg	0.591	0.962	0.604	1.260
Köln	0.741	1.509	0.749	1.418

Table 6. Home and away offensive and defensive strengths in Italy.

Team	Home		Away	
	Offensive Strengths	Defensive Strengths	Offensive Strengths	Defensive Strengths
Juventus	1.652	0.317	1.762	0.582
Napoli	1.575	0.770	1.461	0.400
Roma	1.128	0.816	1.315	0.345
Lazio	1.652	0.816	1.891	1.018
Inter	1.347	0.687	1.179	0.461
Milan	0.768	0.680	1.332	0.945
Atalanta	1.092	0.773	1.225	0.768
Fiorentina	0.983	0.944	1.179	0.729
Sampdoria	1.310	0.858	0.862	1.420
Torino	1.056	0.773	1.043	1.036
Sassuolo	0.423	0.951	0.773	1.345
Genoa	0.845	1.133	0.430	0.582
Bologna	0.910	1.116	0.680	0.959
Udinese	0.884	1.359	1.031	1.200
Chievo	0.807	1.087	0.602	1.272
Cagliari	0.653	1.359	0.645	1.127
SPAL	0.730	1.269	0.730	1.091
Crotone	0.837	1.073	0.726	1.497
Hellas Verona	0.510	1.502	0.680	1.573
Benevento	0.837	1.717	0.454	1.650

Table 7. Home and away offensive and defensive strengths in Spain.

Team	Home		Away	
	Offensive Strengths	Defensive Strengths	Offensive Strengths	Defensive Strengths
Barcelona	1.870	0.536	2.120	0.613
Atletico Madrid	1.007	0.292	1.290	0.476
Real Madrid	1.840	0.923	1.848	0.790
Valencia	1.223	0.730	1.336	0.749
Villarreal	1.187	0.974	1.014	0.953
Real Betis	1.193	1.430	1.119	0.970
Sevilla	1.079	1.071	0.829	1.225
Getafe	0.886	0.600	0.730	0.718
Eibar	0.886	0.877	0.778	1.042
Real Sociedad	1.601	1.338	0.924	1.042
Girona	0.886	1.015	1.070	1.293
Deportivo Alaves	0.716	1.061	0.924	0.934
Celta Vigo	1.079	1.023	1.152	1.259
Espanyol	0.681	0.738	0.730	0.934
Levante	0.852	1.292	0.827	0.934
Athletic Bilbao	0.683	0.730	1.014	1.123
Leganes	0.575	0.828	0.691	1.089
Deportivo La Coruna	0.750	1.522	0.730	1.473
Las Palmas	0.504	1.753	0.415	1.225
Malaga	0.504	1.266	0.461	1.157

Table 8. Home and away offensive and defensive strengths in Turkey.

Team	Home		Away	
	Offensive Strengths	Defensive Strengths	Offensive Strengths	Defensive Strengths
Galatasaray	1.627	0.416	1.376	0.901
Fenerbahçe	1.466	1.030	1.665	0.459
Medipol Başakşehir	1.240	0.589	1.203	0.706
Beşiktaş	1.466	0.441	1.156	0.706
Trabzonspor	1.127	1.374	1.388	0.812
Göztepe	1.127	1.079	0.879	0.954
Sivasspor	0.920	0.923	0.885	1.051
Kasımpaşa	1.132	1.339	1.130	0.976
Kayserispor	0.884	1.200	0.836	0.976
Yeni Malatyaspor	0.827	0.932	0.601	0.848
Bursaspor	0.849	0.831	0.934	1.088
Akhisarspor	0.676	1.128	1.110	1.024
Antalyaspor	0.884	1.062	0.639	1.238
Aytemiz Alanyaspor	1.015	1.275	1.156	1.095
Atiker Konyaspor	0.849	0.693	0.590	0.901
Osmanlıspor FK	0.920	1.108	1.081	1.276
Gençlerbirliği	0.639	0.687	0.879	1.413
Karabükspor	0.354	1.893	0.491	1.576

Table 9. Week actual outcomes versus model outcomes for Premier League from England.

Home Team -Away Team	x-y	Actual Result	Model Result	Result
Huddersfield -Arsenal	0 - 1	Under 2.5 goal	Under 2.5 goal	Correct
West Ham Un -Everton	3 - 1	Over 2.5 goal	Under 2.5 goal	False
Manchester United-Watford	1 - 0	Under 2.5 goal	Over 2.5 goal	False
Southampton-Manchester City	0 - 1	Under 2.5 goal	Over 2.5 goal	False
Newcastle United-Chelsea	3 - 0	Over 2.5 goal	Under 2.5 goal	False
Swansea City-Stoke City	1 - 2	Over 2.5 goal	Under 2.5 goal	False
Liverpool-Brighton & Hove Albion	4 - 0	Over 2.5 goal	Under 2.5 goal	False
Tottenham-Leicester City	5 - 4	Over 2.5 goal	Over 2.5 goal	Correct
Crystal Palace-West Bromwich	2 - 0	Under 2.5 goal	Under 2.5 goal	Correct
Burnley-Bournemouth	1 - 2	Over 2.5 goal	Under 2.5 goal	False

Table 10. Week actual outcomes versus model outcomes for Ligue 1 from France.

Home Team -Away Team	x-y	Actual Result	Model Result	Result
NantesRC-Strasbourg	1 - 0	Under 2.5 goal	Under 2.5 goal	Correct
Metz-Bordeaux	0 - 4	Over 2.5 goal	Over 2.5 goal	Correct
Rennes-Montpellier	1 - 1	Under 2.5 goal	Under 2.5 goal	Correct
Toulouse-Guingamp	2 - 1	Over 2.5 goal	Under 2.5 goal	False
Lyon-Nice	3 - 2	Over 2.5 goal	Over 2.5 goal	Correct
Troyes-Monaco	0 - 3	Over 2.5 goal	Over 2.5 goal	Correct
Marseille-Amiens SC	2 - 1	Over 2.5 goal	Under 2.5 goal	False
Caen-Paris Saint Germain	0 - 0	Under 2.5 goal	Under 2.5 goal	Correct
Saint Etienne-Lille	5 - 0	Over 2.5 goal	Over 2.5 goal	Correct
Dijon-Angers	2 - 1	Over 2.5 goal	Over 2.5 goal	Correct

Table 11. Week actual outcomes versus model outcomes for Bundesliga from Germany.

Home Team -Away Team	x-y	Actual Result	Model Result	Result
Hamburg -Mönchengladbach	2 - 1	Over 2.5 goal	Over 2.5 goal	Correct
Bayern Münch -Stuttgart	1 - 4	Over 2.5 goal	Over 2.5 goal	Correct
Bayer Leverkusen -Hannover 96	3 - 2	Over 2.5 goal	Over 2.5 goal	Correct
Freiburg-Augsburg	2 - 0	Under 2.5 goal	Under 2.5 goal	Correct
Mainz 05-Werder Bremen	1 - 2	Over 2.5 goal	Over 2.5 goal	Correct
Wolfsburg-Köln	4 - 1	Over 2.5 goal	Over 2.5 goal	Correct
Hoffenheim-Borussia Dortmund	3 - 1	Over 2.5 goal	Over 2.5 goal	Correct
Hertha Berlin-RB Leipzig	2 - 6	Over 2.5 goal	Over 2.5 goal	Correct
Schalke 04-Eintracht Frankfurt	1 - 0	Under 2.5 goal	Under 2.5 goal	Correct

Table 12. Week actual outcomes versus model outcomes for Serie A from Italy.

Home Team Away Team	x-y	Actual Result	Model Result	Result
Juventus -Hellas Verona	2 - 1	Over 2.5 goal	Over 2.5 goal	Correct
Genoa -Torino	1 - 2	Over 2.5 goal	Over 2.5 goal	Correct
Napoli -Crotone	2 - 1	Over 2.5 goal	Over 2.5 goal	Correct
Udinese -Bologna	1 - 0	Under 2.5 goal	Under 2.5 goal	Correct
Cagliari -Atalanta	1 - 0	Under 2.5 goal	Over 2.5 goal	False
Chievo -Benevento	1 - 0	Under 2.5 goal	Under 2.5 goal	Correct
SPAL -Sampdoria	3 - 1	Over 2.5 goal	Over 2.5 goal	Correct
Milan -Fiorentina	5 - 1	Over 2.5 goal	Under 2.5 goal	False
Sassuolo -Roma	0 - 1	Under 2.5 goal	Under 2.5 goal	Correct
Lazio -Inter	2 - 3	Over 2.5 goal	Under 2.5 goal	False

Table 13. Week actual outcomes versus model outcomes for La Liga from Spain.

Home Team Away Team	x-y	Actual Result	Model Result	Result
Celta Vigo-Levante	4 - 2	Over 2.5 goal	Under 2.5 goal	False
Leganes-Real Betis	3 - 2	Over 2.5 goal	Under 2.5 goal	False
Malaga-Getafe	0 - 1	Under 2.5 goal	Under 2.5 goal	Correct
Sevilla-Deportivo Alaves	1 - 0	Under 2.5 goal	Over 2.5 goal	False
Las Palmas-Girona	1 - 2	Over 2.5 goal	Over 2.5 goal	Correct
Villarreal-Real Madrid	2 - 2	Over 2.5 goal	Over 2.5 goal	Correct
Valencia-Deportivo La Coruna	2 - 1	Over 2.5 goal	Over 2.5 goal	Correct
Athletic Bilbao-Espanyol	0 - 1	Under 2.5 goal	Under 2.5 goal	Correct
Atletico Madrid-Eibar	2 - 2	Over 2.5 goal	Over 2.5 goal	False
Barcelona-Real Sociedad	1 - 0	Under 2.5 goal	Over 2.5 goal	False

Table 14. Week actual outcomes versus model outcomes for Super League from Turkey.

Home Team- Away Team	x-y	Actual Result	Model Result	Result
Akhisarspor -Osmanlıspor FK	2 - 1	Over 2.5 goal	Over 2.5 goal	Correct
Yeni Malatyaspor -Kayserispor	3 - 2	Over 2.5 goal	Under 2.5 goal	False
Gençlerbirliği -Bursaspor	1 - 0	Under 2.5 goal	Under 2.5 goal	Correct
Aytemiz Alanyaspor- Antalyaspor	3 - 2	Over 2.5 goal	Over 2.5 goal	Correct
Trabzonspor -Karabükspor	3 - 0	Over 2.5 goal	Over 2.5 goal	Correct
Beşiktaş -Sivasspor	5 - 1	Over 2.5 goal	Over 2.5 goal	Correct
Fenerbahçe -Atiker Konyaspor	3 - 2	Over 2.5 goal	Over 2.5 goal	Correct
Göztepe -Galatasaray	0 - 1	Under 2.5 goal	Over 2.5 goal	False
Medipol Başakşehir- Kasımpaşa	3 - 2	Over 2.5 goal	Over 2.5 goal	Correct