



Journal of Human Sport and Exercise

E-ISSN: 1988-5202

jhse@ua.es

Universidad de Alicante

España

RATHKE, ALEX

An examination of expected goals and shot efficiency in soccer

Journal of Human Sport and Exercise, vol. 12, núm. 2, marzo, 2017, pp. 514-529

Universidad de Alicante

Alicante, España

Available in: <http://www.redalyc.org/articulo.oa?id=301052437005>

- How to cite
- Complete issue
- More information about this article
- Journal's homepage in redalyc.org

redalyc.org

Scientific Information System

Network of Scientific Journals from Latin America, the Caribbean, Spain and Portugal

Non-profit academic project, developed under the open access initiative

An examination of expected goals and shot efficiency in soccer

ALEX RATHKE 

Department of Physical Education and Sport Sciences, University of Limerick, Limerick, Ireland

ABSTRACT

The aim of this study was to examine goal scoring in European football leagues and specifically which factors are associated with predicting Expected Goals (xG). This concept helps us to evaluate player, specifically strikers, in the number of goals they score season upon season. Therefore, this study examined the shots from the Premier League and Bundesliga games (380 & 306) from the 2012-2013 season. All of the shots were grouped into sections on the field of play and a theoretical goal value was applied to each area. The factors analysed were: distance of the shot taken from goal and the angle of the shot in relation to the goal. In calculating xG, it is suggested that the distance and angle of the shots matter. A combination of the two factors calculated xG better than each variable alone. Furthermore, this examination of xG has been able to identify mid-table teams scoring and conceding goals relatively accurately. Top league teams and lower league teams over and under achieved respectively. Managers and Coaches may find this method useful in identifying players who consistently score close to their expected total or even out-perform it. **Key words:** BUNDESLIGA, GOAL SCORING, PERFORMANCE ANALYSIS, PLAYERS, PREMIER LEAGUE, FOOTBALL.

Cite this article as:

Rathke, A. (2017). An examination of expected goals and shot efficiency in soccer. *Journal of Human Sport and Exercise*, 12(2proc), S514-S529. doi:<https://doi.org/10.14198/jhse.2017.12.Proc2.05>



Corresponding author. *Department of Physical Education and Sport Sciences, University of Limerick, Limerick, Ireland.*

E-mail: rathkealex@gmail.com

6th ISPAS International Workshop, 22-23 March 2016. International Society of Performance Analysis of Sport. Carlow, Ireland.

JOURNAL OF HUMAN SPORT & EXERCISE ISSN 1988-5202

© Faculty of Education. University of Alicante

doi:10.14198/jhse.2017.12.Proc2.05

INTRODUCTION

Shooting in soccer has been a popular topic for scrutiny among academics and practitioners alike. This is seen through studies by (Armatas et al., 2009; Armatas, 2010; Lago-Peñas et al., 2010; Radman et al., 2014; Szwarc 2004 & Szwarc 2007) to name a few. Hughes & Bartlett (2002) & Szwarc (2007) named goal-scoring as the ultimate measure of performance. A number of other studies stated that winning teams shoot more at goal and therefore are more effective (Bosca et al., 2009; Lago-Peñas et al., 2010; Lago-Peñas et al., 2011; Perin et al., 2013; Szwarc, 2004 & Szwarc 2007). What does effective mean? Does it mean that being able to score more goals sits you higher in a league table? This is the view shared by Bekris et al., (2013) & Mara et al., (2012). However how do these winning teams become more effective at winning and scoring more goals than their opposition? The following paragraph will examine more closely some of the factors that studies in the academic world of football have discussed.

The location of the ball before the shot (Armatas, 2010; Bekris et al., 2013 and Pollard & Reep, 1997), the shot speed (Dörge et al., 2002; Manolopoulos et al., 2006; Potthast et al., 2010 and Radman et al., 2014), the accuracy of the kick/ shots on target (Bar-Eli et al., 2009; Bekris et al., 2013; Castellano et al., 2012; Finnoff et al., 2002; Jankovic et al., 2011; Liu et al., 2015 and Radman et al., 2014) were identified as important criteria for success. Other factors discussed in lesser detail were the time of the game when a goal was scored (Armatas, 2010 and Pratas et al., 2012), the distance away from the goal (Miller & Bartlett, 1996 and Pollard & Reep, 1997), the shot probability (Pollard & Reep, 1997), the possession of the ball (Carmichael et al., 2001 & Lago-Peñas et al., 2011), the player's use of both feet (Carey et al., 2001), the shoes they wore against barefoot shooting (Lees et al., 2010) and the quality of the opposition (Pratas et al., 2012).

The expected goals method

As all these factors have been analysed separately in various studies (as mentioned above), they have not been analysed together in terms of what or which factors give the best indication of scoring the most goals by an individual player. This is what the Expected Goals Method (xG) strives to calculate. In its simplest form, this method involves calculating the chances a team has to score and concede goals. The definition for calculating xG has been interpreted similarly among various blogs and websites such as (Bertin, 2015a; Caley, 2015; Caley, 2014a; Pleuler, 2014; Trainor & Chappas, 2013 and 11tegen11, 2014). The units of measurement are 0 (no goal) to 1 (goal) (Bertin, 2015a and Riley, 2014). Through the research conducted on this topic, this seems to be the first academic study undertaken in professional football. Other sports such as Ice Hockey have used xG to evaluate players as can be seen from Macdonald's (2012) presentation on the topic at the MIT Sloan Conference in 2012. However a number of statisticians or football enthusiasts such as Altman (2015a) Altman (2015b), Bertin (2015a) Bertin (2015b), Bertin (2015c), Caley (2015), Caley (2014a), Eastwood (2014a), Eastwood (2014b), Eastwood (2014c), Gardiner (2015), Pleuler (2014), Riley (2014), Trainor and Chappas (2013) and 11tegen11 (2014) have undertaken studies on player and team xG through personal blogs and websites.

Each of the blogs mentioned have built xG models using different factors such as location, distance, shot speed, shot angle and type of shot (Riley, 2014) with conflicting results (Bertin, 2015a and Pleuler, 2014). Caley has by far done the most research and study on xG which can all be accessed via his blog (Cartilage Free Captain). He however, uses what appears to be a different methodology in every blog that he publishes. In his posts, he uses different zones on the pitch (Caley 2013a; Caley 2014b & Caley 2014e) and also appears to focus only on Shots on Target (SoT) (Caley, 2013a) rather than all shots. Riley (2014) also used only SoT in his blog posts.

Another example of Caley's work can be seen in his 2014a paper where he measured a number of different factors in his study such as location, shot type, speed of attack, passes and set-plays. However his methodology could be questioned. He used a combination of different seasons and included in his method the results table. Interestingly, the inclusion of a club's wage bill as an external factor of performance has been suggested to impact on a team's quality (Caley, 2014d). A completely different view was taken by Bertin (2015c) who reckons that by just using distance as an indicator, a pretty accurate xG method can be calculated. Elsewhere, a more mathematical measure of calculating xG has been examined by Eastwood (2014b) although he also used distance and angle (x and y co-ordinates).

Why is the xG method useful in football? The model has shown to have practical applications in the football world particularly at club level in Denmark. The football club FC Midtjylland won their first Danish league title using this method for the recruitment of players (de Hoog, 2015). While they for obvious reasons have kept their model and calculations a secret, this shows that there is success in providing coaches and clubs with such statistical methods in order to enhance a club's future prospects. Elsewhere, 11tegen11 (2015) state that xG is a good indicator of a team's future performance.

The model has however also received criticisms on a number of levels. Due to the number of factors that need to be included in calculating xG, there is no one model that accurately measures the most significant factors of xG (Bertin, 2015c). Bertin (2015a) critically questions ExG. He mentions that "studies done which looked at only the distance of the shot from goal resulted in a probability rating of (0.997) which is ridiculously high and therefore no other factors need to be factored into the calculation". It has since been examined that there is not necessarily only one factor that gives an accurate evaluation of a team/individual's xG (Pleuler, 2014).

Defensive pressure (Dp) has shown to be another area that has received critical attention in the study on xG (Caley, 2014c; Caley, 2015; Riley, 2014; 11tegen11, 2014; Trainor & Chappas, 2013 & Ward, 2015) and to what extent it reduces the xG value of a shot (Caley, 2014c). The lack of defensive positioning of players in accordance to game specific actions (crosses, through balls and passes) was described by 11tegen11 (2014). A different view of calculating Dp and xG conceded was examined by Ward (2015).

A third criticism of the model is described by Bertin (2015a) and Caley (2015). According to the two, the model does not take into account renowned clubs performances during games and explain why they are better than the rest (Bertin, 2015c & Caley, 2015). Surely though this could be easily explained in the sense that they have better players (while not always efficient) and bigger resources in out-spending their fellow league competitors. Within this a further criticism is mentioned by Caley. The model according to Caley (2015) does not show which player takes the shots, which can be seen as a negative factor. De Hoog (2015) explains however, that through this model, FC Midtjylland were able to identify Finnish midfielder Tim Sparv as their newest recruit.

While the criticism is not without reason, the xG method is still seen as a valuable tool in predicting players and team's goal scoring and conceding probabilities. This was noted by Bertin (2015a) particularly if a model is good and it installs confidence, or if player/team(s) out-perform the model. Central to the xG method is prediction, otherwise noted by Gardiner (2015) as "historical trends". Caley (2015) has already discussed a number of other factors which could affect xG, while others have described them as 'limitless' (Bertin, 2015c; Pleuler, 2014 & Riley, 2014). Therefore, the model is still very much a work in progress (Bertin, 2015a & Caley, 2015). As mentioned above, the need for a more central approach with one model being considered and tested is required (such as the model FC Midtjylland use) as opposed to numerous models.

The paper by Riley (2014) examined only the shots on target along with the shots location. Using only shots on target gives coaches and managers a false indication of a player's xG quality. Therefore this paper will examine the notion that only knowing the shot's distance from goal, xG in soccer can be measured accurately. Furthermore shot specific data will be used in order to calculate team's xG attacking and defensive values.

Following on from this, the study conducted by Trainor & Chappas (2013) examined Goal Expectation and Efficiency in Soccer. They state that shooting efficiency is defined as "the number of actual goals divided by the number of expected goals". Therefore, once the teams' xG has been calculated, their shot efficiency output will be determined. In their study (Trainor & Chappas, 2013) of examining shot efficiency in Europe's Top 5 leagues, they found that a high number of Europe's elite clubs registered a low attacking shot efficiency (Barcelona, Borussia Dortmund and Bayern Munich were the exceptions). It must be remembered that they kept their xG method secret (Bertin, 2015a) and therefore their calculations and predictions may not have been accurate. A contrary view of kicking accuracy is explained by Finnoff and his colleagues (2002). They mention that the amount of shots at the goal and goals scored per match are poor indicators of shot efficiency. They state that (2002:349) "it favours players in positions that shoot on goal more frequently". How this makes it an inaccurate measurement of kicking performance is questionable because that is why there are positions in football with each having different roles, abilities and requirements. For this reason, the second part of this paper will aim to examine the shot efficiency of those teams.

MATERIAL AND METHODS

Sample size

In order to carry out this study, all league games from the English Premier League (380) and the German Bundesliga (306) for the 2012-2013 season were analysed. The collected data set for both leagues was provided by Opta. They are a private company committed to collecting and analysing the performance of teams in most European Leagues (www.optasports.com). Liu et al., (2013) have examined and verified Opta data as a reliable data source. Therefore, we can establish that the previous studies in examining xG (American Soccer Analysis, 2014; Bertin, 2015a; Caley, 2014c; Caley, 2015; Harkins, 2015; Pleuler, 2014 & Riley, 2014) have been conducted using a reliable data source. Additionally, Perin and his colleagues (2013) explain Opta's extensive role in football analysis. After removing penalties from the dataset and not including own goals, the sample size included 18,218 shots. This included missed, saved, blocked and scored shots in both leagues.

Statistical analysis

A logistic regression was performed on the dataset using SPSS (Version 22). SPSS processed that all shots within the dataset were accounted for but that they would miss the target. When the distance and the angle of the shots taken were added, no superior result was produced. A closer look at the dataset revealed no definite issues. A lack of variables which were not included in the dataset were free-kicks, through balls and corner kicks. These were suspected to produce the issue. The variable factor of penalty kicks was included in the dataset, however these were removed for the purpose of the study. The three variables (free-kicks, through balls and corner kicks) are important components in creating a model on which a statistical test such as logistic regression can be based. Therefore, based on the data provided, the author maintains that a statistical test such a logistic regression could not be performed.

The hypothesis of this paper was stated as 'examining the notion that only knowing the shot's distance from goal, xG in soccer can be measured accurately'. Instead of providing no statistical measurement to test the hypothesis, the author provided the breakdown of all shots and goals from each zone (Table 1). In the section

(Calculating Expected Goals) below, the calculations (Table 1) were used to test the hypothesis to an acceptable extent. For the record, 52 shots (Premier League) and 29 shots (Bundesliga) as well as own goals were not included in this study (Table 1). These shots were taken in the 'other half' of the pitch and not deemed suitable for calculating ExG.

Calculating expected goals

The xG model was built by dividing all shots into eight zones using an x (distance from goal) and y (angle from goal) co-ordinate as shown in (Figure 1) and by the recommended guidelines of Taylor (2014). The zones 1-5 are all located in the eighteen-yard box, with zones 6 and 8 further out and zones 7 covering the sides of the pitch. The original model was taken from Caley (2013a, 2014b & 2014e) and modified to map all areas of the pitch. Once all shots were distributed to a zone, the amount of shots taken per zone, the SoT and goals in that zone were calculated and recorded (these values are known as the league average numbers). In addition, the number of goals from total shots per zone and shots on target per zone were calculated as a percentage (Table 1). These percentages were considered as the xG value (probability) that shots were converted from those zones.

The same procedure was used to calculate each clubs' shots per zone, SoT and number of goals scored. Furthermore, the number of shots against them from each zone and goals scored from each zone were recorded. In order to calculate each clubs xG value for the season, their shots per zone numbers were multiplied by its corresponding %goals per shots number.

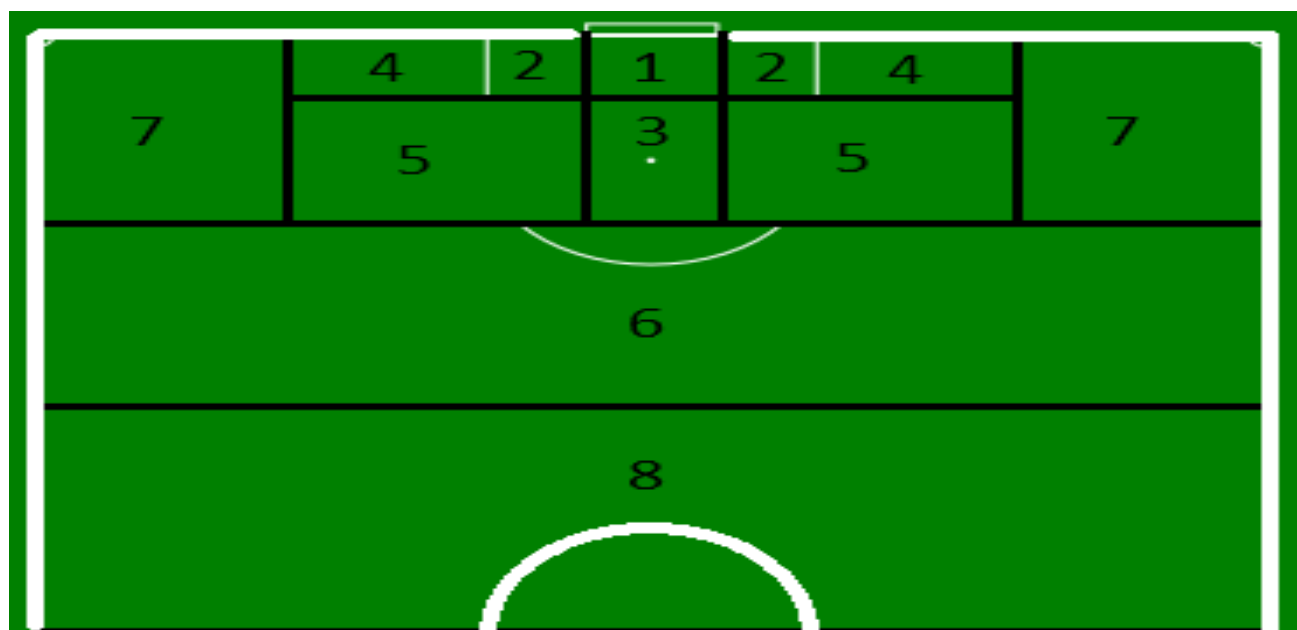


Figure 1. Pitch zones.

Shot efficiency

The statement "the number of actual goals divided by the number of expected goals" mentioned above by Trainor & Chappas (2013) was used to measure shot efficiency for this paper.

Ethics

Ethical approval was not required for this paper. However a confidentiality contract was signed with Opta for the use of the data in this study.

RESULTS

Statistical results

Taking only distance into account from the frame of the goal (Zone 1 & 3), these shots were worth more goals than from other zones (Premier League Z1: 34% & Z3: 15%; Bundesliga Z1: 38% & Z3: 17%). The shots taken from directly in front of goal revealed a higher expectation of scoring than those at an angle. Nevertheless, distance and angle can both be seen to have an impact on calculating xG. In addition, shots from inside the box (Zones 1-5) were converted more than shots outside of the box. While there is no data to suggest how all of these shots were delivered into the different zones, an observational difference nonetheless can be observed. Interestingly, the SoT from all zones stay reasonably constant (Premier League: Z1= 66%; Z2= 59%; Z3= 60%; Z4= 63%; Z5= 65%; Z6= 59%; Z7= 57% and Z8= 54%) (Bundesliga: Z1= 67%; Z2= 64%; Z3= 59%; Z4= 65%; Z5= 65%; Z6= 59%; Z7= 55 and Z8= 49%).

Premier League	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6	Zone 7	Zone 8	Total
No. of shots	631	113	2,822	183	2,111	3,657	35	929	10,481
SoT	418	67	1,706	116	1,373	2,171	20	501	5,742
%SoT	66%	59%	60%	63%	65%	59%	57%	54%	
Goals per zone	214	11	415	13	143	136	4	23	959
%G from shots per zone	34%	10%	15%	7%	7%	4%	11%	2%	
%G from SoT	51%	16%	24%	11%	10%	6%	20%	5%	
Bundesliga	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6	Zone 7	Zone 8	Total
No. of shots	427	80	2,084	155	1,743	2,311	56	880	7,737
SoT	287	51	1,226	101	1,126	1,354	31	435	4,611
%SoT	67%	64%	59%	65%	65%	59%	55%	49%	
Goals per zone	161	10	362	18	151	97	2	17	818
%G from shots per zone	38%	13%	17%	12%	9%	4%	4%	2%	
%G from SoT	56%	20%	30%	18%	13%	7%	6%	4%	
Both League Total									18,218

Figure 2. Zonal Calculations.

Bundesliga results



Figure 3. Bundesliga team's xG vs actual goals scored.

xG vs Actual goals per game

Bayern Munich and Borussia Dortmund ended the 2012/13 season ahead of the rest in terms of scoring a higher number of goals per game (*Munich*: 2.76 & *Dortmund*: 2.26) than they were expected to score (*Munich*: 2.04 & *Dortmund*: 1.79) given the chances they had. In addition, *Mainz 05*, *Borussia Mönchengladbach*, *Fortuna Düsseldorf*, *Hannover 96* and *Schalke* scored a higher number of goals per game (1.12, 1.24, 1.06, 1.62 & 1.68) than they were expected to (1.04, 1.24, 0.90, 1.26 & 1.61). Meanwhile, *Augsburg*, *Freiburg*, *Greuther Fürth*, *Werder Bremen*, *Hoffenheim* and *Stuttgart* underachieved in their scoring expectation (1.29, 1.34, 1.07, 1.48, 1.15 & 1.21) to (0.97, 1.21, 0.74, 1.35, 1.09 & 1.00).

According to this model, the Bundesliga had a higher number of clubs which conceded more goals than they were expected to concede than their Premier League counterparts (*Eintracht Frankfurt*= 1.32:1.21; *Augsburg*= 1.41:1.24; *Schalke*= 1.35:1.09; *Fortuna Düsseldorf*= 1.62:1.46; *Hannover 96*= 1.79:1.49; *Hoffenheim*= 1.91:1.19; *Greuther Fürth*= 1.65:1.36; *Stuttgart*= 1.56:1.37; *Werder Bremen*= 1.82:1.13 and *Wolfsburg*= 1.38:1.13). Bayern Munich, Bayer Leverkusen, Borussia Mönchengladbach, Freiburg and Nürnberg were the exceptions (0.47, 1.00, 1.26, 1.06 & 1.26) to (0.66, 1.12, 1.35, 1.16 & 1.28).

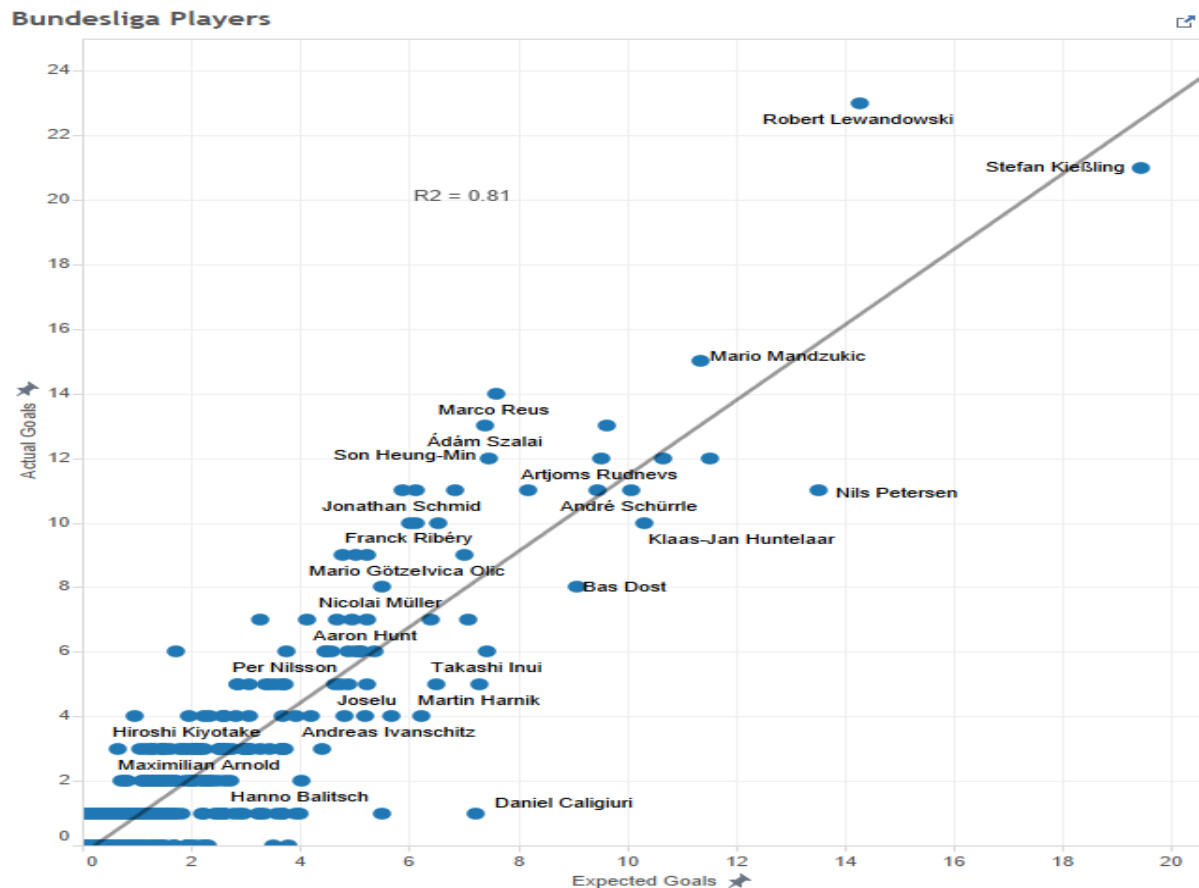


Figure 4. Bundesliga Player's xG vs Actual Goals Scored.

Bundesliga Players

We can see that the relationship between xG and Goals scored is quite high (0.81). If we look at the individual players in that season, we notice a big gap between the first two players and the rest. This does not mean that in this case Lewandowski and Kießling are better than the rest. It purely means that they a) either took more shots than the other players or b) that these shots were of a higher quality (in this case closer to goal). Examining the data, both players were in the Top 3 for most shots taken (98 and 120) that season and took 69% and 78% of their shots in the penalty area. In this case, therefore, we can say that while they did take more shots, these were generally of a higher quality.

Premier league results

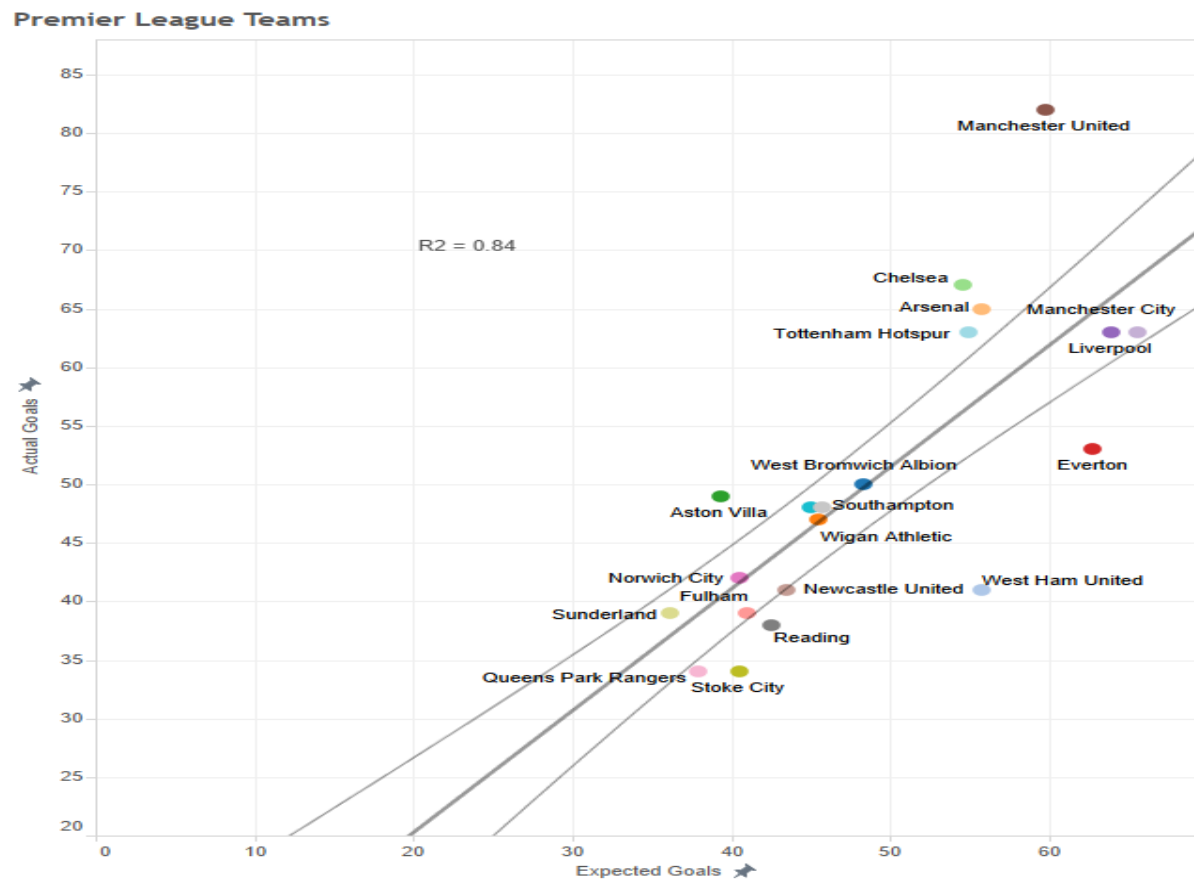


Figure 5. Premier League Team's xG vs Actual Goals Scored.

xG vs Actual goals per game

The champions *Manchester United* had the highest expected versus actual scoring per game difference (1.57 to 2.21) of all Premier League teams. Other teams to score a greater number of goals than what was expected of them were: *Aston Villa* (1.16:1.03), *Arsenal* (1.79:1.47), *Chelsea* (1.79:1.43), *Liverpool* (1.74:1.68), & *Tottenham* (1.74:1.44). In comparison, *Everton*, *Stoke City*, *Queens Park Rangers* and *West Ham United* greatly underachieved in their scoring expectation (1.65, 1.07, 1.00 & 1.47) to (1.39, 0.82, 0.79 & 1.11).

In terms of defensive xG metrics, *Sunderland* were the only Premier League side to concede considerably less goals (1.32) than they were expected to (1.48). Given the chances, *Aston Villa* (1.68:1.41), *Liverpool* (1.11:0.87), *Newcastle United* (1.76:1.27), *Norwich City* (1.45:1.12), *Reading* (1.92:1.67), *Southampton* (1.45:1.05), *Tottenham Hotspurs* (1.18:0.85) and *Wigan Athletic* (1.84:1.22) conceded a greater number of goals than they were expected to.

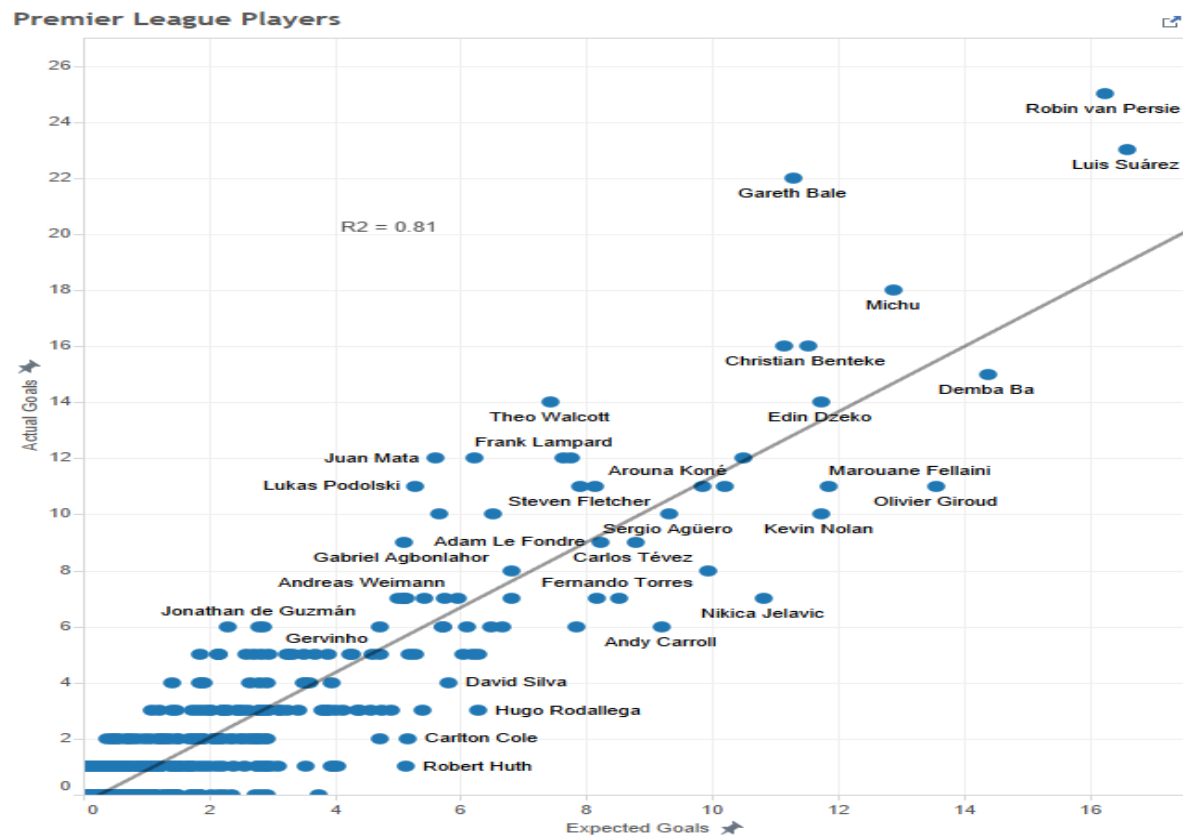


Figure 6. Premier League Player's xG vs Actual Goals.

Premier League Players

The same method can be used to explain the gap (smaller than the Bundesliga) between Bale, Van Persie & Suarez. The three players took the most shots that season (165, 141 and 187) of which 38%, 62% and 59% came in the penalty area. The smaller percentage number for Bale can be attributed to his position (attacking midfielder – while the other two are strikers) that season and having the responsibility of hitting free-kicks for Tottenham.

Shot efficiency

The Top Ten teams rated for Overall Shot Efficiency were:

1. Bayern Munich (1.92)	2. Manchester United (1.44)
3. Borussia Mönchengladbach (1.27)	4. Chelsea (1.26)
5. Borussia Dortmund (1.19)	6. Sunderland (1.15)
7. Arsenal (1.11)	8. Bayer 04 Leverkusen (1.09)
9. Hannover 96 (1.07)	10. Fortuna Düsseldorf (1.05) / West Brom (1.05)

Notable absentees from this list include Hamburg (ranked 17th with 0.97), Manchester City (ranked 18th with 0.97), Tottenham Hotspurs (ranked 21st with 0.86), Wolfsburg (ranked 22nd with 0.86), Liverpool (ranked with 26th with 0.81) & Stuttgart (ranked 31st with 0.72). This model suggests that the three worst efficient teams in this study, originate from the Bundesliga (Hoffenheim: 36th with 0.58, Werder Bremen: 37th with 0.57 & Greuther Fürth: 38th with 0.56).

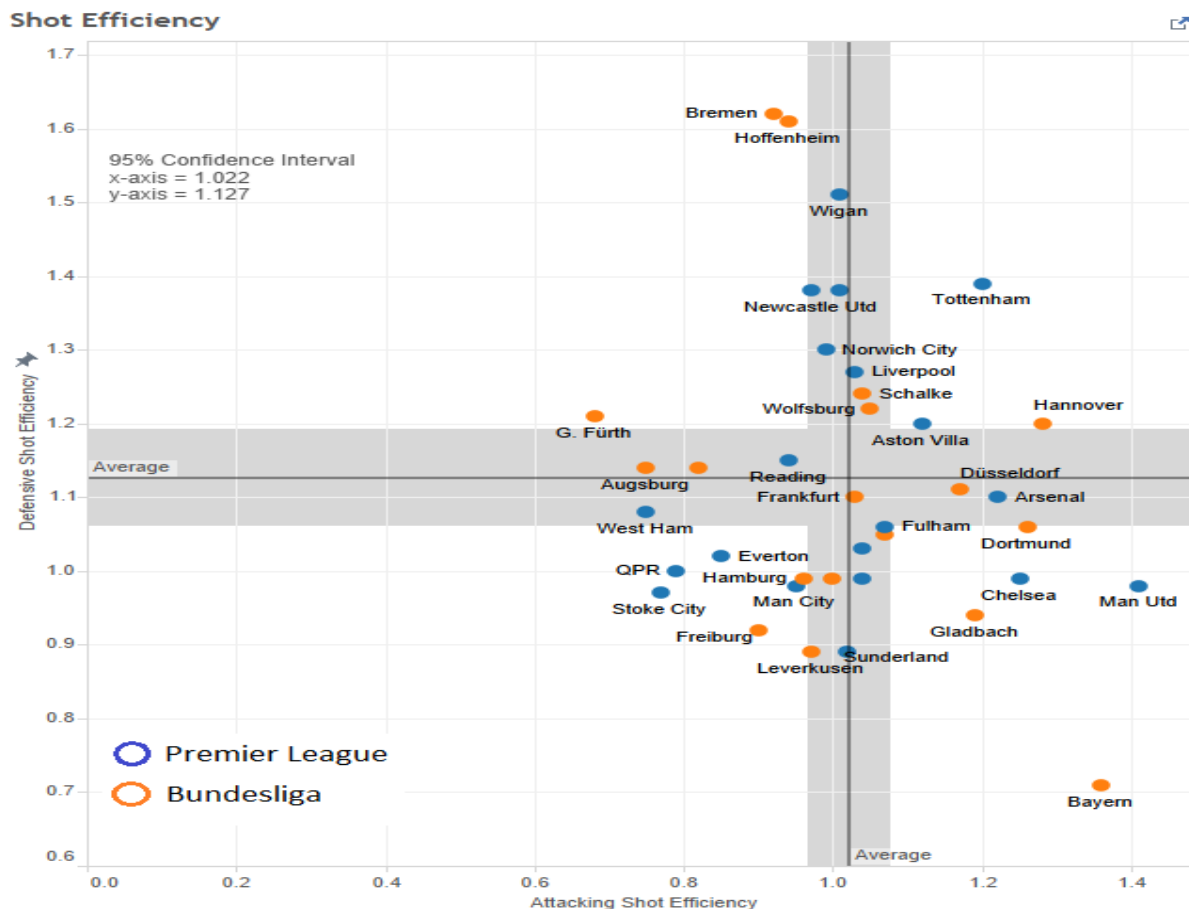


Figure 7. Shot Efficiency of Premier League & Bundesliga Teams.

DISCUSSION

The aim of this study was to 1) identify if distance alone can accurately measure xG in soccer and 2) complete a study on xG and Shot Efficiency in soccer. Although the first objective of the study may be common sense to many soccer players, coaches and fans, it is still important to observe the different aspects of what makes a good expected goal count. The second part of this study hopes to contribute to the knowledge and analysis of football which could be used productively in training sessions for tactical studies. Similarly, it could also help clubs to identify players and pursue them through the transfer market.

The results of this study suggest that distance does matter when calculating xG. However, a combination of the distance and the angle of the shot from goal may be better suited to calculating xG. Herewith, this paper rejects the hypothesis, which was stated as 'examining the notion that only knowing the shot's distance from goal, xG in soccer can be measured accurately'. The results of this paper contradict the views of Bertin (2015c). Similar results were seen through Eastwood's (2014a) blog post of adding the Y factor (angle) to the X factor (distance).

Further in-depth study of statistical analysis within xG is required. This is to ensure that a correct statistical procedure is applied, in determining which variables are more significant to consider than others. Within calculating xG, there may only be a small number of variables that predict xG more accurately than the measures described by Bertin (2015c), Pleuler (2014) and Riley (2014).

Further results of this study indicate that the method of calculating xG can be considered reliable when measuring clubs' expected performances for the current and future seasons. Caley (2013a), Eastwood (2014b) and Riley (2014) have published similar results yet the numerical differences for each club could be explained through the methodology used. A separate website (The Chels) which published Riley's (2014) work has also found similar results.

Football results are difficult to predict correctly and xG is no exception. The number of variables that can be counted (playing styles, free-kicks, headers, pass types and dribbles) are endless and will all have some impact on a player's xG total (some more than others). Similar views are shared by Caley (2013b) and Caley (2013c). It is also worth noting that each club's stadium pitch sizes are different and this factor may also impact the precise positioning of a shot (Parton, 2013).

In another instance, xG does not measure well against the top teams of a league as they often are able to score a higher amount of goals based on their team's quality. This view is shared in studies of xG by Bertin (2015c), Caley (2015) & Eastwood (2014b). This idea may be explained alone in the previous paragraph, however a lot of debate has focused on the lack of defensive data (Bertin, 2015a; Caley, 2015 & Tegen, 2014) available in xG. This view is not shared by Ward (2015). In the meantime, results from a study conducted by Harkins (2015) examined how the defensive metrics in basketball could be applied to football.

The results of shot efficiency studies in this paper are also dependent on the above outlined factors. Not as much research has been undertaken on shot efficiency as there has been for xG. While Trainor & Chappas (2013) methodology is unknown, this paper also ranks Bayern Munich, Sunderland and Borussia Mönchengladbach in the Top 10.

As discussed throughout this paper, the topic of xG is relatively new within the football community and consequently, limited research has been undertaken. For instance, Eastwood (2014c) examined the difference of headed vs footed shots and their xG value. Elsewhere, comparisons on xG values by shot and pass types have been studied in European leagues by Caley (2014e) & Caley (2014f), while Altman (2015b) built his own xG model (non-based shots) based on using passes instead of shots. Herewith, this paper suggests to not rely too heavily on professional club football examples to aid decisions at International or Youth football. It is also suggested that further research is conducted on the topic of xG. This paper outlines five areas for further research, which include 1) statistical evaluation of xG variables, such as recently conducted by American Soccer Analysis (2017) 2) if xG should be applied at International matches 3) if playing styles and tactics influence xG, 4) if wages and transfer activity impact on a club's xG and 5) if xG in women's football differs to men's football.

CONCLUSIONS

This study has demonstrated the value and reliability that xG has within professional football. The variables of distance and angle together were seen to have a major impact of calculating xG rather than distance as a variable alone. There may be no direct practical application of this method available yet, however it could be incorporated into training exercises (attacking and defensive) to aid player's understanding and needs of the game. For example, how attacking players strike certain shots and how defenders should be positioned to defend these shots.

From previously established research in the field, we know that xG can help predict future performance. Therefore, this method could help clubs identify potential transfer targets and statistically evaluate them.

However, a precaution which needs to be considered is that this method has not been examined in more detail as of yet regarding the different types of shot types, passes, through balls and set-plays.

ACKNOWLEDGEMENTS

This research would not have been possible without the data, which was provided by Opta. The author would also like to thank Mr Rob Carroll (Performance Analysis Lecturer at the University of Limerick) for his most valuable insight and for sharing his knowledge throughout the course of this research. The author would also like to thank his family and friends for the support provided throughout this time. Any errors are my own and should not cause damage to the respective individuals' reputations.

REFERENCES

1. American Soccer Analysis. (2014). Expected Goals 2 Explanation. *American Soccer Analysis*. Retrieved from <http://www.americansocceranalysis.com/expected-goals-2-explanation/>
2. American Soccer Analysis. (2017). Validating the ASA xGoals Model. *American Soccer Analysis*. Retrieved from <http://www.americansocceranalysis.com/home/2017/3/6/validating-the-asa-xgoals-model>
3. Altman, D. (2015a). Pitfalls of measuring shooting and saving skill. *North Yard Analytics Blog*. Retrieved from <http://www.northyrdanalytics.com/blog/2015/08/22/pitfalls-of-measuring-shooting-and-saving-skill/>
4. Altman, D. (2015b). Beyond Shots: OptaPro Forum 2015. *North Yard Analytics Blog*. Retrieved from <http://www.northyrdanalytics.com/blog/2015/02/05/beyond-shots-optapro-forum-2015/>
5. Armatas, V., Yiannakos, A., Papadopoulou, S., & Skoufas, D. (2009). Evaluation of goals scored in top ranking soccer matches: Greek "Superleague" 2006-07. *Serbian Journal of Sports Sciences*, 3(1), 39-43.
6. Armatas, V. & Yiannakos, A. (2010). Analysis and evaluation of goals scored in 2006 World Cup. *Journal of Sport and Health Research*, 2(2), 119-128.
7. Bar-Eli, M. & Azar, O. (2009). Penalty kicks in soccer: an empirical analysis of shooting strategies and goalkeepers' preferences. *Soccer & Society*, 10(2), 183-191.
8. Bekris, E., Mylonis, E., Sarakinos, A., Gissis, I., Gioldasis, A. & Sotiropoulos, A. (2013). Offense and defense statistical indicators that determine the Greek Superleague teams and placement on the table 2011-2012. *Journal of Physical Education and Sport*, 13(3), 338-347.
9. Bertin, M. (2015a). Why Soccer's Most Popular Advanced Stat Kind Of Sucks. *Regression*. Retrieved from <http://regression.deadspin.com/why-soccers-most-popular-advanced-stat-kind-of-sucks-1685563075>
10. Bertin, M. (2015b). EXPG Comment Reply. *Michael Bertin Blog*. Retrieved from <http://www.michaelbertin.com/expg-comment-reply/>
11. Bertin, M. (2015c). The third to last thing I'll ever write about expected goals. *Michael Bertin Blog*. Retrieved from <http://www.michaelbertin.com/the-third-to-last-thing-ill-ever-write-about-expected-goals/>
12. Boscá, J., Liern, V., Martinez, A. & Sala, R. (2009). Increasing offensive or defensive efficiency? An analysis of Italian and Spanish football. *Omega*, 37, 63-78.
13. Caley, M. (2013a). Shot Matrix 1: Shot Location and Expected Goals. *Cartilage Free Captain*. Retrieved from <http://cartilagefreecaptain.sbnation.com/2013/11/13/5098186/shot-matrix-i-shot-location-and-expected-goals>

14. Caley, M. (2013b). Shot Matrix III: The Incredible Through-Ball. *Cartilage Free Captain*. Retrieved from <http://cartilagefreecaptain.sbnation.com/2013/11/16/5111212/shot-matrix-iii-the-incredible-through-ball>
15. Caley, M. (2013c). Shot Matrix II: Pass Type and Shot Type, or, Heading is super hard. *Cartilage Free Captain*. Retrieved from <http://cartilagefreecaptain.sbnation.com/2013/11/15/5107438/shot-matrix-ii-pass-type-and-shot-type-or-heading-is-super-hard>
16. Caley, M. (2014a). Premier League projections, from the winners to the relegated clubs. *Cartilage Free Captain*. Retrieved from <http://cartilagefreecaptain.sbnation.com/2014/9/11/6131661/premier-league-projections-2014#methodology>
17. Caley, M. (2014b). English Premier League Advanced Statistics. *Cartilage Free Captain*. Retrieved from <http://cartilagefreecaptain.sbnation.com/2014/2/12/5404348/english-premier-league-shot-statistics>
18. Caley, M. (2014c). Pre-shot dribbling in soccer is extremely valuable on scoring chances. *Washington Post*. Retrieved from <https://www.washingtonpost.com/news/fancy-stats/wp/2014/08/27/pre-shot-dribbling-in-soccer-is-extremely-valuable-on-scoring-chances/>
19. Caley, M. (2014d). Here are the Premier League teams that have made the most of their payrolls. *Washington Post*. Retrieved from <https://www.washingtonpost.com/news/fancy-stats/wp/2014/09/05/here-are-the-premier-league-teams-that-have-made-the-most-of-their-payrolls/>
20. Caley, M. (2014e). Shot Matrix International 1: Shot Distribution in European Football. *Cartilage Free Captain*. Retrieved from <https://cartilagefreecaptain.sbnation.com/2014/2/4/5375492/shot-matrix-international-i-shot-distribution-in-european-football>
21. Caley, M. (2014f). Shot Matrix International II: Shot Type and Pass Type. *Cartilage Free Captain*. Retrieved from <http://cartilagefreecaptain.sbnation.com/2014/2/5/5379690/shot-matrix-international-ii-shot-type-and-pass-type>
22. Caley, M. (2015). Let's talk about expected goals. *Cartilage Free Captain*. Retrieved from <http://cartilagefreecaptain.sbnation.com/2015/4/10/8381071/football-statistics-expected-goals-michael-caley-deadspin>
23. Carey, D., Smith, G., Smith, D., Shepard, J., Skriver, J., Ord, L. & Rutland, A. (2001). Footedness in world soccer: an analysis of France '98'. *Journal of Sports Sciences*, 19(11), 855-864.
24. Carmichael, F., Thomas, D. & Ward, R. (2001). Production and Efficiency in Association Football. *Journal of Sports Economics*, 2(3), 228-243.
25. Castellano, J., Casamichana, D. & Lago, C. (2012). The use of match statistics that discriminate between successful and unsuccessful soccer teams. *Journal of Human Kinetics*, 31, 139-147.
26. De Hoog, M. (2015). How data, not people, call the shots in Denmark. *The Correspondent*. Retrieved 24 September 2015, from <https://decorrespondent.nl/2607/How-data-not-humans-run-this-Danish-football-club/517995289284-77644562>
27. Dörge, H., Andersen, T., Sørensen, H. & Simonsen, E. (2002). Biomechanical differences in soccer kicking with the preferred and the non-preferred leg. *Journal of Sports Sciences*, 20(4), 293-299.
28. Eastwood, M. (2014a). Expected Goals: The Y Axis. *Pena.It/y Blog*. Retrieved from <http://www.pena.it/y/2014/04/16/expected-goals-the-y-axis/>
29. Eastwood, M. (2014b). Expected Goals for all. *Pena.It/y Blog*. Retrieved from <http://www.pena.it/y/2014/02/12/expected-goals-for-all/>
30. Eastwood, M. (2014c). Expected Goals: Foot Shots versus Headers. *Pena.It/y Blog*. Retrieved from <http://www.pena.it/y/2014/08/28/expected-goals-foot-shots-versus-headers/>
31. Finnoff, J. T., Newcomer, K. & Laskowski, E. R. (2002). A valid and reliable method for measuring the kicking accuracy of soccer players. *Journal of Science and Medicine in Sport*, 5(4), 348-353.

32. Gardiner, B. (2015). Were Swansea Lucky to Finish 8th?. *StatsBomb*. Retrieved from <http://statsbomb.com/2015/06/were-swansea-lucky-to-finish-8th/>
33. Harkins, J. (2015). Assessing defensive impact. *OptaPro Blog*. Retrieved from <http://www.optasportspro.com/about/optapro-blog/posts/2015/blog-assessing-defensive-impact/>
34. Jankovic, A., Leontijevic, B., Pasic, M. & Jelusic, V. (2011). Influence of certain tactical attacking patterns on the result achieved by teams participants of the 2010 FIFA World Cup in South Africa. *Physical Culture*, 65(1), 34-45.
35. 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 & Medicine*, 9, 288-293.
36. Lago-Peñas, C., Lago-Ballesteros, J. & Rey, E. (2011). Differences in performance indicators between winning and losing teams in the UEFA Champions League. *Journal of Human Kinetics*, 27, 135-146.
37. Lees, A., Asai, T., Andersen, T., Nunome, H. & Sterzing, T. (2010). The biomechanics of kicking in soccer: A review. *Journal of Sports Sciences*, 28(8), 805-817.
38. Liu, H., Hopkins, W., Gomez, A. & Molinuevo, J. (2013). Inter-operator reliability of live football match statistics from OPTA Sportsdata. *International Journal of Performance Analysis in Sport*, 13, 803-821.
39. Liu, H., Gomez, M., Lago-Peñas, C. & Sampaio, J. (2015). Match statistics related to winning in the group stage of the 2014 Brazil FIFA World Cup. *Journal of Sports Sciences*, 33(12), 1205-1213.
40. Macdonald, B. (2012). An Expected Goals Model for Evaluating NHL Teams and Players. In *MIT Sloan Sports Analytics Conference* (pp. 1-8). New York: United States Military Academy. Retrieved from <http://www.sloansportsconference.com/wp-content/uploads/2012/02/NHL-Expected-Goals-Brian-Macdonald.pdf>
41. Mara, J., Wheeler, K. & Lyons, K. (2012). Attacking Strategies that lead to goal scoring opportunities in high level women's football. *International Journal of Sports Science & Coaching*, 7(3), 565-577.
42. Manolopoulos, E., Papadopoulos, C. & Kellis, E. (2006). Effects of combined strength and kick coordination training on soccer kick biomechanics in amateur players. *Scandinavian Journal of Medicine in Science of Sports*, 16, 102-110.
43. Parton, S. (2013). 2013-2014 Premiership Pitch Sizes. *Openplay*. Retrieved from <http://www.openplay.co.uk/blog/premiership-football-pitch-sizes-2013-2014/>
44. Perin, C., Vuillemot, R. & Fekete, J. (2013). Soccer Stories: A Kick-off for Visual Soccer Analysis. *Visualisation and Computer Graphics, IEEE Transactions*, 19(2), 2506-2513.
45. Pleuler, D. (2014). On the topic of Expected Goals and the repeatability of finishing skill. *OptaPro Blog*. Retrieved from <http://www.optasportspro.com/about/optapro-blog/posts/2014/on-the-topic-of-expected-goals-and-the-repeatability-of-finishing-skill.aspx>
46. Potthast, W. & Brüggemann, G. (2010). Motion-differences in goal kicking on natural and artificial soccer turf systems. *Footwear Science* 2.
47. Pratas, J., Volossovitch, A. & Ferreira, A. (2012). The effect of situational variables on teams' performance in offensive sequences ending in a shot on goal. A Case Study. *The Open Sports Sciences Journal*, 5, 193-199.
48. Radman, I., Wessner, B., Ruzic, I., Markovic, G., Hackl, M., & Bachl, N. (2014). Influences of graded physiological exertion on shooting ability in soccer. In *World Congress of Performance Analysis of Sport X*. Opatija: Faculty of Kinesiology, University of Zagreb, Croatia.
49. Riley, P. (2014). A shooting Model – An Exp(G)lanation and Application. *Differentgame*. Retrieved from <https://differentgame.wordpress.com/2014/05/19/a-shooting-model-an-expgplanation-and-application/>

50. Swarc, A. (2004). Effectiveness of Brazilian and German Teams and the teams defeated by them during the 17th FIFA World Cup. *Kinesiology*, 36(1), 83-89.
51. Swarc, A. (2007). Efficacy of Successful and Unsuccessful Soccer Teams Taking Part in Finals of Champions League, *MEDSPORTPRESS*, 13(2), 221-225.
52. Taylor, M. (2014). How to calculate expected goal model. *Pinnacle Sports*. Retrieved 13 October 2015, from <http://www.pinnaclesports.com/en/betting-articles/soccer/how-to-calculate-goal-models>
53. TheChels. (2010). Expected Goal Statistics. *TheChels.co.uk*. Retrieved 18 November 2015, from <https://www.thechels.co.uk/expected-goals/>
54. Trainor, C. & Chappas, C. (2013). Goal Expectation and Efficiency. *StatsBomb*. Retrieved from <http://statsbomb.com/2013/08/08/goal-expectation-and-efficiency/>
55. Ward, D. (2015). How To Measure Defense? *Statsbomb*. Retrieved from <http://statsbomb.com/2015/11/how-to-measure-defense/>
56. 11tegen11. (2014). Expected Goals 2.0 – Some light in the black box. *11tegen11*. Retrieved from <http://11tegen11.net/2014/08/07/expected-goals-2-0-some-light-in-the-black-box/>
57. 11tegen11. (2015). The best predictor for future performance is Expected Goals, *11tegen11*. Retrieved from <http://11tegen11.net/2015/01/05/the-best-predictor-for-future-performance-is-expected-goals/>