

Pattern Analysis of Goals Scored in Malaysia Super League 2015

Nazim Razali¹, Aida Mustapha², Filipe Manuel Clemente³, Md Fauzi Ahmad⁴,
Mohamad Aizi Salamat⁵

^{1,2,5}Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia

⁴Faculty of Business and Technology Management, Universiti Tun Hussein Onn Malaysia

³School of Sport and Leisure, Viana do Castelo Polytechnic Institute, Portugal

Article Info

Article history:

Received Feb 12, 2018

Revised Apr 20, 2018

Accepted Apr 27, 2018

Keywords:

Association football

Soccer

Football analysis

Goals scored analysis

Malaysian Super League

ABSTRACT

This paper is set to study the pattern of goals scored across 132 matches in 2015 Malaysian Super League (MSL). Despite of well-known researches in European football, no information in Malaysian football teams has been done consistently. To identify the patterns, this research focuses on goals scored based on timing that includes scoring frequency, the impact of first goal achieved by team towards the final match outcome, as well as the impact of having home advantage for matches during MSL 2015. The findings of this paper will provide useful information on MSL besides helping coaches with fresh insight for creating effective training and tactical plan as the time progressed during match. At the same time, players will also be more prepared for a consistent performance especially when training for matches as the visiting teams.

Copyright © 2018 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Nazim Razali,

Faculty of Computer Science and Information Technology,

Universiti Tun Hussein Onn Malaysia.

Email: aidam@uthm.edu.my

1. INTRODUCTION

In association football (British English) or soccer (American English), data from match statistics usually provide records of meaningful events that occur during the match, such as the number goals scored with time, name of scorer, name of yellow and red card receivers or name of injured players. Match statistics also include raw technical data such as the number of shots fired on and off target, total number of passes or number of foul committed. After every match, the data is then processed and analyzed to form an analysis report that can guide both coaches and players in assessing the team performance during the match. Many researchers have used information on past ranking in the final league, goals difference, total goals scored by the team, goals conceded by the team and final points gained end of the season to conduct a match analysis. At present, match analysis in association football covers analysis on game flow [1, 2] team performance [3, 4] or individual player performance [5, 6].

In addition, match statistics are also used for football prediction modelling and development, in particular for predicting match outcome based on past-match statistical data [7, 8], used the points gained by the team from five previous seasons and current season for developing the first phase of their prediction model. This model is called the PI football, which focuses on team strengths [7]. Meanwhile, [8] used the past results of match data to run their bivariate Poisson model for match prediction. Following [7] and [8], this paper is set to analyze the pattern of goals scored in 2015 Malaysia Super League (MSL) matches by focusing on the relationship between goals scoring patterns and the period of match time. The analysis will cover 410 goals across 132 matches scored by 12 teams, whereby each team played 22 matches including home and away.

The remainder of this paper proceeds as follows. Section 2 reviews previous works on match analysis, Section 3 presents the proposed methodology for pattern analysis in MSL, Section 4 presents and discusses the results, and finally Section 5 concludes this work with some indication for future research.

2. RELATED WORK

In 2014, an intensive and systematic review on match analysis in football was carried by [2]. The scope of the reviews was divided into three categories, which included descriptive analysis (focusing on description of players activity patterns), comparative analysis (focusing on relationship between parameters that may influence the player performance such as player position, game location and fatigue level), and predictive analysis (focusing on the effective ways of playing football which may influence the outcomes of matches). Meanwhile, [1] focused on analyzing ball possession during the Spanish football league in 2008-2009. The research examined all matches in terms of match results such as win, draw and losses, as well as match location and quality of opposition before relating all the factors back to ball possession in proportion of match time.

A notational analysis was carried out in order to study team performance, which may provide useful information to the player and coach by describing the relevant parameter consisting of defensive and offensive strategies. Both strategies are very important to increase the team efficacy [3, 4]. Notational analysis is applicable on the team and match-play sports, whereby it is important to study the relationships between the movement and behaviour of individual team players. Similar analysis was carried out by Duch et al. [6] to measuring and quantifying both teams and players performance as inspired by methods in social network analysis. In subsequent research, [5] developed a model called the EA Sports Player Performance Index that was used to measure and quantify each individual player performance during the English Premier League (EPL).

Interestingly, analysis of the goal patterns scored in football matches has also been widely studied, whereby the primary objective is to examine and evaluate the pattern of goals scored at certain football competition. [9] carried out their studies on evaluating the number of goals scored pattern in Greek Super League 2006-2007 while [10] and [11] analyzed the goal scoring pattern during the European Championship 2012. Later, [12] expanded the scale of studies by analyzing goal scoring patterns from top four European Leagues National Championships, which include the English Premier League, the French Football Ligue 1, the Italian Serie A, and the Spanish Football Liga for three seasons (2008-2009, 2009-2010 and 2010-2011). Despite of well-known researches in European football, no information in Malaysian football teams have been done consistently. Based on that, the aim of this study is to identify the patterns of goals scored in 132 games from the Malaysia Super League.

3. METHODS

The central focus of this study is to analyze the goal scoring patterns in the Malaysian Super League (MSL), which is a top-tier professional football league in Malaysia league system since the establishment. Although this league is categorized as grade D football league, there are rapid revolutions undergoing the Malaysian football leagues recently that warrant for further improvements.

All the datasets used for this analysis including match fixtures, goal scored, time during goal was scored, and venue of matches were extracted from the official website of Football Association of Malaysia (FAM) (<http://www.fam.org.my/>) through its Football Informational Management System (<https://www.pengurusanbolasepakfam.org.my/>). All the data was cleaned and prepared using the IBM SPSS Statistics version 22 and R version 3.3.3 with “ez” package.

This study examined all the match statistics individually, with a total of 410 goals scored in 132 matches from 12 teams, both performing home and away. The study then adopted the techniques used in [9], with exception of using the aid from video analysis program and assistant from experienced observers due to unavailability of such video data. Next, the Generalized Eta Squared (η_G^2) from the work of [13] is used to measure the effect size of goals scored frequency per intervals of 45 minutes and 15 minutes. The elements that will be used to analyze scoring patterns during the Malaysia Super Leagues (MSL) season 2015 include:

- Goals scored frequency per 45 minutes intervals (first half including extra time and second half including extra time).
- Goals scored per 15 minutes intervals (1-15 minutes, 16-30 minutes, 31-45 minutes + extra time, 46-60 minutes, 61-75 minutes and 76-90 + extra time).
- First goals impacts on the match outcome for the scoring team (lose, draw and win).
- Home advantage in tournament (away win, draw and home win).

In representing the results, descriptive statistics consisting of frequency distribution were used. The statistical data were reproduced with absolute frequency for the number of goals scored and the number of matches, as well as the relative frequency for the number of goals scored in percentage. A non-parametric chi-square analysis was carried out to assess the statistical significance of differences in goals scored within the frequency of every 45-minute period and 15-minute period, together with the impact of first goal towards match outcome as well as the home advantage. In addition, the Generalized Eta Squared (n_G^2) is also measured using repeated measures of ANOVA for identifying the effect size of goals between periods of the matches. [14] stated that measuring the effect size is important in order to observe the association or differentiation between variables in formal experimental research. According to [15], the Generalized Eta Squared (n_G^2) is more suitable as compared to Partial Eta Squared (n_p^2) and Eta Squared (n^2) for providing comparison between effect sizes. The values of effect size are measured by 0.2; minimum, 0.5; moderate and 0.8; strong and the level of significance was set at $p < 0.05$.

4. RESULTS AND DISCUSSION

From the 410 goals scored during the 2015 Malaysia Super League (MSL), the frequency of goals scored during the first half period and the second half period is shown in Figure 1.

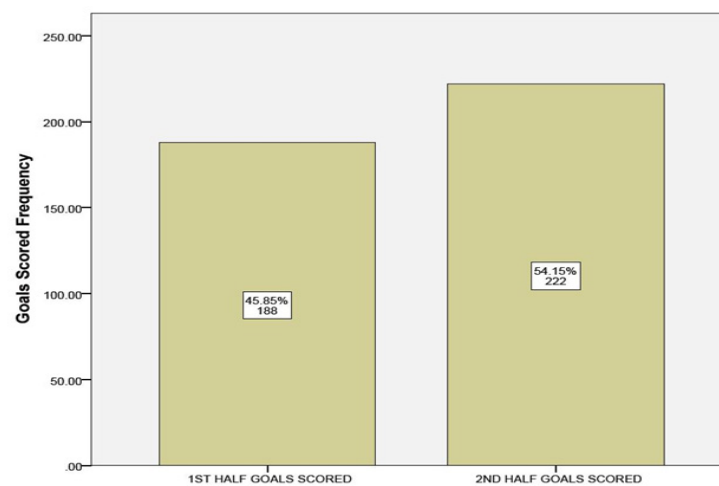


Figure 1. Goals Scored Frequency per 45 minutes

From Figure 1, although the teams scored more goals during the second half period (45.85%) than the first half period (54.15%), no statistical differences were observed in chi-squared analysis ($p > 0.05$). However, repeated measures test shown that the minimum effect size for first half period is $n_G^2 = 0.053$ with $p < 0.05$, while the minimum effect size for second half period $n_G^2 = 0.032$ with $p < 0.05$.

The results are in line with related studies on the pattern of goals scored, whereby the frequency of goals scored is dependent on the period of match time. For example, [9] revealed that there was a significant increase in the frequency of goals scored as the match time progressed during the Greek Super League season 2006-2007. Similar results were found on the European Champion 2012 [10,11] as well as four major European Leagues (English Premier League, French Football Ligue 1, Italian Serie A and Spanish Football Liga) across three consecutive seasons which are 2008-2009, 2009-2010 and 2010-2011 [12]. All the works indicated that the frequency of goals scored in second half period (46-90 minutes + extra time) is significantly higher than the first half period (1-45 minutes + extra time) of match. However, the frequency of goals studied by [10] and [11] did not show steady increment as compared to [9] and [12] for interval of 15 minutes in a football match.

Next, Figure 2 shows that the chi-square analysis revealed a significant difference in goals scored for every 15 minutes ($p < 0.001$) and the repeated measures showed results of the 45 minutes intervals which are as follow: first 15 minutes ($p > 0.05$; $n_G^2 = 0.020$, minimum effect size), 16-30 minutes ($p > 0.05$; $n_G^2 = 0.0002$, minimum effect size), 31-45 minutes ($p < 0.05$; $n_G^2 = 0.0613$, minimum effect size), 46-60 minutes ($p > 0.05$; $n_G^2 = 0.009$, minimum effect size), 61-75 minutes ($p < 0.05$; $n_G^2 = 0.043$, minimum effect size) and final 15 minutes interval ($p > 0.05$; $n_G^2 = 0.001$, minimum effect size). There are similar percentages of goals

scored in period of 31-45 minutes, 61-75 minutes and 76-90 minutes with additional extra time. The highest percentage of goals scored achieved in 46-60 minutes while the lowest percentages of goals scored in 16-30 minutes. However, the repeated measures showed minimum effect size for 15-interval period of time.

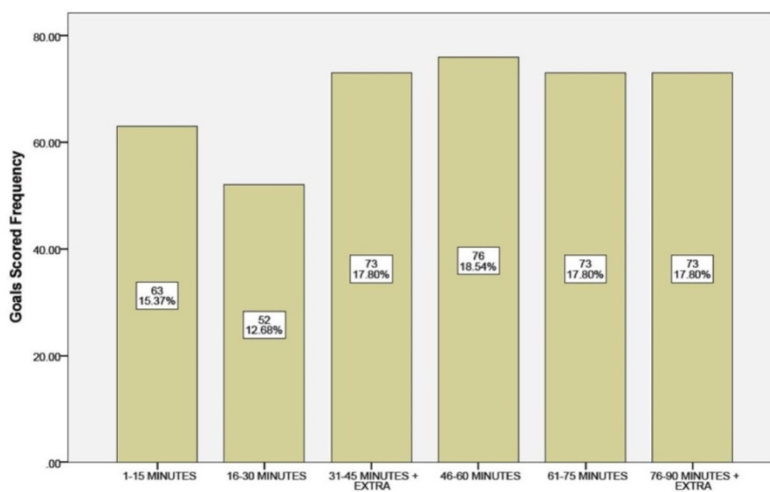


Figure 2. Goals Scored Frequency per 15 minutes

Research in [9-12] agreed that the frequency of goals scored is higher during the second half of a match despite fatigue among the players. Following the findings, this research assumed that the frequency of goals scored in Malaysia Super League 2015 is high during the last 15 minutes before half time as well as during the second half period of the match due to fatigueness.

The development of fatigue during match may lead to deteriorate of physical performance, technical performance, tactical effectiveness and mental condition. According to [16], distance covered by sprinting and most intense running during 15 minutes interval period for football players keep decreasing as the time progressed. Besides, high intensity running and sprinting during the final 15 minutes showed that substitute player covered more distance as compared to players who played the entire match.

Fatigueness is a major concern during football matches because it may lead to areas not being covered by the team during defensive, hence causing more open space to the opposition team. In addition, a newly substituted player may cause problem to the opponent's player whom is playing from the beginning of the match. This in turn, will increase the chances and opportunities to score goals. [17] stated that distance covered and work rates are influenced by the players' position and roles. As the result, midfield players covered more distance in total as compared to the defenders and forwards. However, [9] stated that defenders are in the position that are prone to fatigueness, then may cause lapse in concentration hence giving advantage to the attacking team.

In research on effect of playing formation towards high intensity running and technical profiles in English FA Premier League, [18] found that midfielders covered greater distance in high intensity running alongside the attackers as compared to defenders. All of this may lead to development of fatigue for individual player according to position and formation. For example, the workload for lone striker in formation 4-5-1 is higher if compared to formation 4-3-3, therefore the players easily experience fatigue. Yet, the workload of defenders is high in formation 4-3-3 as compared to 4-5-1, hence opening them to risk of fatigues. [19] also observed that the players who experience fatigue will have a decrease in involvement with the ball as well as a decrease in the effectiveness of technical performance. This is reflected by lower number of passes and successful passes. Special cases where fatigue does not inflict the performance of players may due to the importance of match like winning the championship as well as avoiding the relegation.

In this work, in order to study the impact of the first goal on match outcomes, five matches were removed from the dataset because there were no goals scored. The matches include KL FELDA United vs. Kelantan, KL FELDA United vs. Lions XII, Lions XII vs. Johor DT, Perak vs. Kelantan and Sarawak vs. Pahang). The results showed a significant difference between win, draw and lose for the first goal scored on the match outcomes ($p < 0.001$). Figure 3 shows that the team that scored the first goal will win the match by 65.35% chances as compared to draw (18.11%) and lose (16.54%).

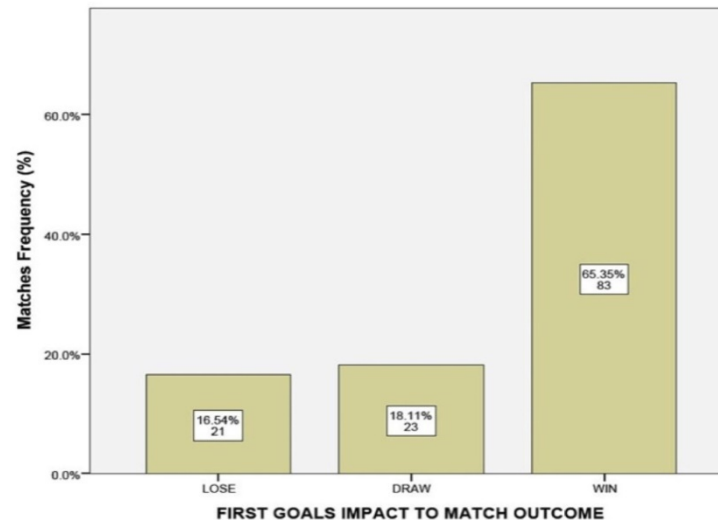


Figure 3. First Goal Impact on Match Outcome

Given the analysis on impact for first goals scored to match outcomes in the MSL, it is assumed that the team who scored first will manage to win the match. [9] able to achieve 71.43% of percentage for win team that scored the first goal in Greek Super League 2006-2007 which then they spotted that the final standings of the Greek Super League showed imbalanced points and goals achieved by top teams. The studies on first goals scored by [10] and [20] supported the findings on the importance of scoring the first goal for the team to increase the probability in winning the match. Meanwhile, [20] conducted a more specific research in first goal scoring for five major European Championship 2014-2015. The research found the relationship between scoring the first goal with having home advantage, whereby 84.85% of home team scored first with home advantage while the away achieved the scoring first advantages about 76.25%. [10] explained that the team who achieved first goal in match will perform better due to increase in confidence level.

Pratas JM [21] used football matches in the Portuguese Premier League 2009-2010 season in order to determine the performance indicators that influence the time when the first goal is scored. In the study, about 12% of matches from overall matches per season have been removed because the matches are ended goal-less. They used performance indicators such as goal difference, shots on goal, sanctions and substitutions in order to predict the time of first goals achieved by home team. The higher the goals difference, the higher the chances for the team to obtain first goal. As the result, they found that home team gained percentage of first goal scored by 57.5%, which resulting about 74.59% win outcome, while away team gained 42.46% for first goal scored and 62.22% resulting to win outcome of the matches. The total percentage of win for overall for teams successfully scored the first goal is 70%. Hence, scoring first goal give big influence to determine the course of match outcome as the time progressed.

Finally, the chi-squared analysis showed that there are significant differences between home win as compared to draw and away win for home advantage ($p < 0.001$). Thus, the home teams have more advantage during match as shown by Figure 4, whereby the percentage of win was 53.79% at home as compared to draw (21.21%) and lose (25%).

The Figure 4 showed a considerably low result in terms of percentage of home advantage if compared to other works. The research on home advantage by [20] revealed that the home advantage for five major European Leagues in season 2014/2015 (English Premier League, French Ligue 1, Spanish La Liga, Italian Seria A and German Bundesliga) are higher than the Malaysian Super League 2015 by 58.58%, 59.98%, 57.58%, 56.47%, and 61.84% respectively. The reason underlying the low percentage for home advantage in MSL is due to the fact that the four teams competing in the MSL used the same venue as home. The teams that share the same venue are Selangor and PDRM (Shah Alam Stadium) as well as KL Sime Darby and KL FELDA United (Selayang Stadium). The match results showed that PDRM only draw (home) and lose (away) to Selangor at the Shah Alam Stadium. The same fact went to KL Sime Darby, whereby they only draw (home) and lose (away) to KL FELDA United at the Selayang Stadium.

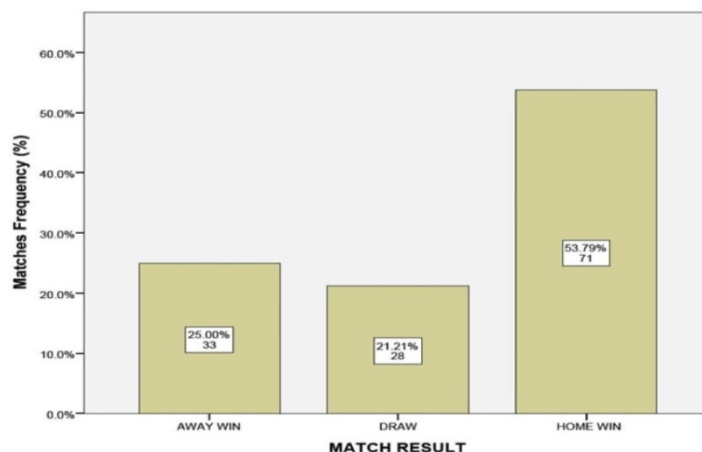


Figure 4. Home Advantage

Another possibility for the low percentage of home advantage is attributed by football teams who are based on organizations or government bodies (KL FELDA United, ATM, PDRM, KL Sime Darby). These teams have lower pool of fans as compared to the state teams in Malaysia. This is because majority of football teams in Malaysia are established and managed by the state government or government-linked company with exception of Johor Darul Takzim (JDT) club, which is owned by HRH Brigadier General Tunku Ismail Ibni Sultan Ibrahim, the State Crown Prince. Also, there is an invited football team from the Republic of Singapore (Lions XII) which has been competing in MSL since 2012.

Further research need to be carried out in order to prove the assumption by [22]. [22] examined nine seasons of game by game performance data in Greek Super League consisting of 2,160 matches in order to evaluate the significance difference between teams in the home advantage. [23] pointed out that home teams performed more successful passes, successful tackles, successful crosses, and successful aerial challenges. A greater of frequency of corner kicks, crosses, dribbles and shots on goals were observed at home teams while visiting team perform more goals kicks, interceptions, clearance and losses possession. Thus, further studies for home advantage are needed towards team performance as it may influence the technical performance, tactical effectiveness and psychology to home team and visiting team considering the importance of match, fan attendance as well as distance travelled.

5. CONCLUSION

The aim of this paper was to quantify and analyze the pattern of goals scored in Malaysia Super League (MSL) for seasons in 2015 since there are no information in Malaysian football teams about this have been done compared to other international football league such as English Premier League, French Ligue 1, Spanish La Liga, Italian Seria A and German Bundesliga. Considering the relationship between goals scoring patterns and the period of match time, the results revealed that the frequency of goals scored increased as the time progressed. The results from the experiments showed that 54.15% goals were scored in second half period of match. Moreover, studies on interval of 15 minutes of time during match showed that the frequency of goals scored is very high during the last 15 minutes of the first half period and the first 15 minutes of the second half period in a match. However, the repeated measures showed there are minimum effect size of frequency of goals scored for both 45 minutes interval period and 15 minutes interval period. The home advantage and first goal impact are also proven to influence the results of football matches by achieved rate of win 53.79% and 65.35%.

Adding additional aid from video analysis for further research may give more information about pattern of goals such as the position in field where the ball is kicked to the goal have be done, goal event detection [24], finding suitable kicking range for goal conversion [25] or techniques on how football player kicked the ball to the goal [26]. The findings of this study provided useful information to coaches and players competing in the Malaysian Super League. The coaches can design more effective training sessions especially on fitness in order improve their team stamina, hence avoiding the team from facing exhaustion towards the end of matches. In addition, the insights from the findings can be used to craft tactical strategies based on match data and capability of the players as a team. This will benefit the players by making them more adaptable to changes in tactical strategies employed by coaches. Moreover, training on psychology is also important to increase the confidence level among the player, hence increasing the psychological

resistance level towards stress and tension during match, especially when come as visiting team and left behind in goal scored.

ACKNOWLEDEMENTS

This research is funded by the Ministry of Higher Education Malaysia under the Research Acculturation Collaborative Effort (RACE) grant scheme Vot 1513.

REFERENCES

- [1] Lago-Peñas C, Dellal A. Ball Possession Strategies in Elite Soccer According to the Evolution of the Match-Score: the Influence of Situational Variables. *Journal of Human Kinetics*. 2010; 25(July): 93–100.
- [2] Sarmento H, Marcelino R, Anguera T, Campanico J, Matos N, Leitao JS. Match analysis in football: a systematic review. *Journal of Sports Sciences*. 2014; 32(20): 1831–1843.
- [3] Clemente F. Study of successful soccer teams on FIFA World Cup 2010. *Pamukkale Journal of Sport Sciences*. 2012; 3(3): 90–103.
- [4] Clemente F, Couceiro M, Martins F, Mendes R. Team's performance on FIFA U17 World Cup 2011: Study based on notational analysis. *Journal of Physical Education and Sport*. 2012; 12(1): 13–17.
- [5] McHale IG, Scarf PA, Folker DE. On the development of a soccer player performance rating system for the English Premier League. *Interfaces*. 2012; 42(4): 339–351.
- [6] Duch J, Waitzman JS, Amaral LAN. Quantifying the performance of individual players in a team activity. *PLoS ONE*. 2010; 5(6): 1–7.
- [7] Constantinou AC, Fenton NE, Neil M. *Pi-football: A Bayesian network model for forecastin Association Football match outcomes*. *Knowledge-Based Systems*. 2012; 36: 322–339.
- [8] Koopman SJ, Lit R. A dynamic bivariate Poisson model for analysing and forecasting match results in the English Premier League. *Journal of the Royal Statistical Society Series A: Statistics in Society*. 2015; 178(1): 167–186.
- [9] Armatas V, Yiannakos A, Papadopoulou S. Evaluation of goals scored in top rankings soccer matches: Greek Superleague 2006-07. *Serbian Journal of Sports Sciences*. 2009; 3(1): 39–43.
- [10] Leite WSS. Euro 2012: Analysis and Evaluation of Goals Scored. *International Journal of Sports Science*. 2013; 3(4): 102–106.
- [11] Michailidis Y, Michailidis C, Primpa E. Analysis of goals scored in European Championship 2012. *Journal of Human Sport and Exercise*. 2012; 8(2): 367–375.
- [12] Alberti G, Iaia MF, Arcelli E, Cavaggioni L, Rampinini E. Goal scoring patterns in major European soccer leagues. *Sport Sciences for Health*. 2013; 9(3): 151–153.
- [13] Bakerman R. Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*. 2015; 37(3): 379–384.
- [14] Richardson J. Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*. 2011; 6(2): 135–147.
- [15] Fritz C, Morris P, Richler J. Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*. 2012; 141(1): 2–18.
- [16] Mohr M, Krstrup P, Bangsbo J. Fatigue in soccer: a brief review. *Journal of sports sciences*. 2005; 23(6): 593–599.
- [17] Di Salvo V, Baron R, Tschann H, Montero F, Bachl N, Pigozzi F. *Performance characteristics according to playing position in elite soccer*. *International Journal of Sports Medicine*. 2007; 28(3): 222–227.
- [18] Bradley PS, Di Mascio M, Peart D, Olsen P, Sheldon B. High Intensity Activity Profiles of Elite Soccer Players at Different Performance Levels. *Journal of Strength and Conditioning Research*. 2010; 24(9): 2343–2351.
- [19] Rampinini E, Impellizzeri FM, Castagna C, Coutts AJ, Wisloff U. Technical performance during soccer matches of the Italian Serie A league: Effect of fatigue and competitive level. *Journal of Science and Medicine in Sport*. 2009; 12(1): 227–233.
- [20] Lago-Peñas C, Gómez-ruano M, Megías-navarro D. Home advantage in football: Examining the effect of scoring first on match outcome in the five major European leagues. *International Journal of Performance Analysis in Sport*. 2016; 16(2): 411–421.
- [21] Pratas JM, Volossovitch A, Carita AI. The effect of performance indicators on the time the first goal is scored in football matches. *International Journal of Performance Analysis in Sport*. 2016; 16(1): 347–354.
- [22] Armatas V, Pollard R. Home advantage in Greek football. *European Journal of Sport Science*. 2014; 14(2): 116–122.
- [23] Tucker W, Mellalieu SD, James N, Taylor JB. Game Location Effects in Professional Soccer: A Case Study. *International Journal of Performance Analysis in Sport*. 2005; 5(2): 23–35.
- [24] Bayat F, Moin M S, Bayat F. Goal Detection in Soccer Video: Role-Based Events Detection Approach. *International Journal of Electrical and Computer Engineering*. 2014; 4(6): 979-988
- [25] Masaleno A, Hassan M M, Muslihudin M, Susilowati T. Finding kicking range of sepak takraw game: Fuzzy logic and Dempster-Shafer theory approach. *Indonesia Journal of Electrical Engineering and Computer Science*. 2016; 2(1): 187-193.
- [26] Rusydi M I, Huda S, Rusydi F, Sucipto M H, Sasaki M. Pattern recognition of overhead forehand and backhand in badminton based on the sign of local Euler angle. *Indonesia Journal of Electrical Engineering and Computer Science*. 2016; 2(3): 625-635.