Benchmarking the Performance of Southeast Asian Football Teams Using the CCR Data Envelopment Analysis (CCR-DEA) Model

Rene D. Estember, Shean Michael L. Reyes, and Omar A. Solaiman
School of Industrial Engineering and Engineering Management
Mapua University, Manila, Philippines
rdestember@mapua.edu.ph, reyes.shean.michael@gmail.com, omarsolaiman028@gmail.com

Abstract

Football is played by approximately 250 million people, and draws approximately 1.3 billion in audiences worldwide, making it the most popular sport in the world. Being a competitive sport, football teams are naturally interested in improving the performance of its players, improving its standing in their respective regions, and gaining a competitive edge over opponents. This paper analyzed the offensive and defensive efficiencies of Southeast Asian football teams that competed in the AFF Suzuki Cup 2018. This was achieved by mathematical optimization method, in particular Data Envelopment Analysis (DEA), to compute for the frontiers of offensive and defensive efficiency. The results from DEA yielded some interesting findings. First, the researchers found that on average, the offensive efficiency of the teams is greater than their defensive efficiency. Second, the researchers also found the teams to be more efficient at home than when playing away. Third, the researchers found that the best strategy to produce points and goals for SEA football teams in the AFF is to improve the defense when playing at home, and to improve the offense when playing away from home.

Keywords
Football, OR in sports, Data envelopment analysis, Performance management, Technical efficiency

1. Introduction

Football is one of, if not the most popular sport in the world. In 2018, Nielsen published a report on football statistics in the days leading up to the FIFA world cup. Football drew far bigger numbers in terms of popularity and audiences, with more than 40% of the population claiming to be interested in the sport (Lovett, 2018). On the other hand, the football's governing body, the *Fédération Internationale de Football Association* (FIFA), estimated that at the turn of the 21st century there were approximately 250 million football players and over 1.3 billion people "interested" in football. In 2010, a combined television audience of more than 26 billion watched football's premier tournament, the quadrennial month-long World Cup finals (Alegi, Joy et al, 2018).

The governing body of football in Southeast Asia is the Association of Southeast Asian Nations (ASEAN) Football Federation or more commonly known as the AFF. The AFF hosts its Championships (otherwise known and from this point will be referred to as the Suzuki Cup) biennially since 1996. There are three stages in the AFF: the qualification round or play-offs, the group competitions or stages, and the semi-finals and finals (ASEAN Football Federation, 2018). In the group stages, the teams are split into two groups, wherein at the end of the group stage, the two top-performing teams from each group advance to the semis and the finals. Based on the AFF tournament regulations, the scoring system is as follows: three (3) points are awarded to the winner of each match, one (1) point if the match concludes in a draw, and none (0) for the defeated team.

In football, even one goal can spell the difference between victory or defeat. Team coaches employ different tactics that they feel will best utilize the abilities of the team's players (Zambom-Ferraresi, Rios, et al, 2018). A study by Zambom-Ferraresi et al (2018) names player selection and design of strategy as the biggest decision factors or variables. Another study by Villa and Lozano (2016) stressed the importance of football teams using their resources effectively to secure a win.

Data Envelopment Analysis or DEA is a non-parametric, linear programming technique that is used to measure the efficiency of Decision-Making Units (DMUs) (Cooper, Seiford, et al, 2004). Vincova (2005) defines DEA models as composed of an *n* number of DMUs utilizing *i* inputs to produce *j* outputs. There are different types of efficiencies measured in DEA based on the model being used. The most common is technical efficiency. Technical efficiency can be defined as the utilization of inputs to transform them into outputs. There are two distinct types of technical efficiency – the overall technical efficiency (OTE) and pure technical efficiency (PTE), and a third called the scale efficiency (SE) (Kumar and Gulati, 2008). The overall technical efficiency falls under the CCR model, while the PTE is under the BCC model. The scale efficiency is the ratio of OTE and PTE and represents the ability of a firm to choose an optimal size or scale of operations to achieve the expected production. Most applications of data envelopment analysis (DEA) in football focus on team performance across games. The structured mechanics of a competitive sport like football make it possible to study the game as a productive activity (Bosca, Liern, et al, 2009).

For example, the study of Villa and Lozano (2016) proposed a network CCR-DEA model that utilized both defensive and attacking variables as resources (inputs) and goals scored as integer output variables. The researchers incorporated the competitiveness of football by studying each match separately and analyzing the efficiency between the two teams parallelly.

DEA has also been applied in sports other than football. One study analyzed major league baseball teams spanning multiple decades (Lewis and Lock, 2009). In another study of Ruiz, Sirvent, et al (2009), DEA has been compared against a weighted scalar index for player assessment in basketball. The researchers found that the results are similar or even advantageous when using DEA. A study on cycling teams in the Tour de France demonstrated DEA benchmarking by evaluating the team as a part of a same group and another evaluation as part of all the teams regardless of the type (Rogge, Puyenbroek, et al, 2012). Other studies have been successful in applying DEA in tennis (Chitnis and Vaidya, 2014), and baseball (Wu, Hsiao, et al, 2015), respectively.

In many competitive sports, the presence of the so-called "home advantage effect" has been extensively studied especially in sports psychology. Citing an earlier study, Anderson et al. defined the home advantage effect as *"consistent finding that home teams win over 50% of the games played under a balanced home and away schedule* (Anderson, Wolfson, et al, 2011). In one study, Nevill, Balmer, et al (1999) studied the influence of the crowd on the decisions of the referee in association football. In this study, the researchers found that officials make more subjective refereeing decisions in favor of the team playing at home, and this effect is especially prominent if the crowd is bigger.

This paper aims to determine the overall, offensive, and defensive efficiency of Southeast Asian football teams that competed in the AFF Suzuki Cup 2018 using DEA and benchmark the football teams to discover the best-performing teams and identify gaps in output for the inefficient teams. The researchers also aim to identify the relationship between input and output variables and how they interact with and affect each other and recommend an optimum strategy for offense and defense when playing as home and away.

2. Methodology

This study adopted a similar research design to Bosca et al (2009) which involves identification of decision-making units (DMUs) to be subjected to DEA. Match statistics that constitute input and output variables were then selected and the data for these variables collected from football publications and websites. An output-oriented CCR-DEA model was used to compute for the efficiency of each DMU. The input variables are also subjected to correlation and regression analysis. Figure 1 summarizes the conceptual design of this paper.

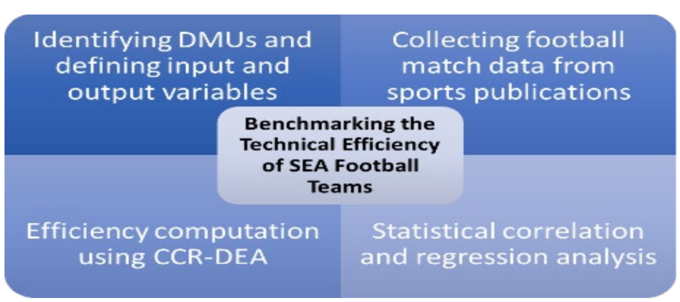

Figure 1. Conceptual Design

2.1 Identification of DMUs

To conduct DEA, the researchers first identified the decision-making units (DMU) to be studied. DMUs are homogeneous entities engaged in the same activities. In this study, each of the participating teams was considered a production unit or a DMU. More specifically, as production DMUs which converts resources into tangible outputs. The DMUs in this paper are the competing teams in the AFF Suzuki Cup 2018, as show in Table 1.

Table 1. Decision-Making Units (DMUs)

DMU No.	Country	DMU No.	Country
1	Vietnam	6	Thailand
2	Malaysia	7	Philippines
3	Myanmar	8	Singapore
4	Cambodia	9	Indonesia
5	Laos	10	East Timor

2.2 Identification of Input and Output Variables

As with any other competitive game, match statistics are also measured in addition to the goals scored by the team. In this study, the researchers utilized these as input variables. Match statistics generally measure two types of abilities of the team: offensive and defensive abilities. Because of this, the researchers also chose to categorize the abilities of a team into offensive inputs and defensive inputs, similar to the method used by Bosca et al (2009). These statistics are utilized as performance indicators that highlight the strengths and weaknesses of each team. In every game, the possession of the ball by each team is measured as well. Figure 2 illustrates the typical Input – Process – Output (I–P–O) of a football team in a match.

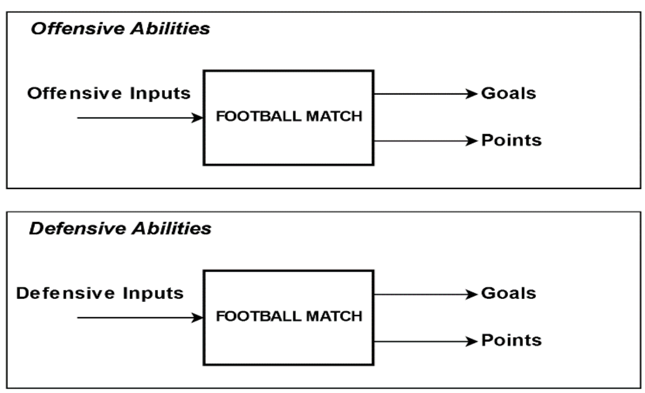

Figure 2. Basic I–P–O of a Football Team in a Match

The offensive inputs of the teams determine their ability to score a goal in each match. In this case, identifying how each of the inputs must be utilized in offensive chances created in attempt to score goals and win games. There are 9 offensive inputs considered, namely: ball possession, shot accuracy, pass accuracy, goal attempts, shots on goal, free kick, corner kick, attacks, and dangerous attacks.

Effective defense is also a key in scoring goals and winning games. As a counterpart to offensive inputs, there are four defensive inputs, namely: saves by goalkeeper, blocked shots, corners defended and fouls.

Every team in the competition aims to qualify for the semi-finals or the championship game. In the AFF Suzuki Cup 2018, 10 teams were divided into 2 groups (Group A: DMUs 1 up to 5; Group B: DMUs 6 up to 10). The two best-performing teams in each group advanced to the semi-finals. The outcome of the tournament is decided upon two outputs: points and goals. Two teams with the highest Points based on the tournament's pointing system will qualify in the semi-finals round of the tournament. As per AFF Suzuki Cup regulations, the winning team gets 3 points regardless of the amount of goals. One point for each team is awarded whenever the outcome is a draw and no point are awarded to the losing team.

2.3 Data Collection

In a football match, match statistics are manually collected and logged by specialized companies entirely devoted to scrutinizing every second of a football match (Bialik, 2014). These match data are eventually broadcasted and make their way to sports stats websites. Therefore, the published football scores one can find on the internet are more or less exactly the same across these sports publications and websites. This eliminates the need to watch each match, so the researchers simply opted to collect match data from these websites and Internet publications. The AFF publishes its match statistics for the biennial AFF Suzuki Cup on its website, but is limited to the current edition, 2018. The AFF website also offers a play-by-play account of the match down to the minute, which formed the basis for the input variable *defended corners.* To gather data for the input variables goal attempts, attacks, dangerous attacks, the researchers gathered data from FlashScore.com, an internet repository of football scores and results for thousands of leagues.

2.4 CCR-DEA Model

The data envelopment model was first introduced by Charnes, Cooper, and Rhode (1978). This model is otherwise known as the CCR model, which can be input or output oriented. The researchers chose an output-oriented model to find out whether the DMUs are producing the maximum output (goals and points) that they can produce given their level of input. Originally, the CCR model is a linear divisive model which can be converted into a traditional linear programming (LP) form [20] as shown below.

$$
\begin{aligned}
&\max \phi \\
&s.t. \sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{io} \\
&\sum_{j=1}^{n} \lambda_j y_{rj} \geq \phi y_{ro} \\
&\lambda_j \geq 0
\end{aligned}
$$

The optimal value of the LP is represented by Φ, but this is not the value of the efficiency. To compute for the technical efficiency, the formula is TE = 1/Φ, meaning a value of Φ = 1 means that a DMU is efficient and a Φ > 1 implies that the DMU is not maximizing their output given their input levels and thus, inefficient. λ represents the unknown weights assigned to each DMU j for input x or output y. The RHS of the model represents the input and output values for the DMU being evaluated.

2.5 Efficiency calculations using CCR-DEA

Three efficiency calculations were made considering all matches: the *overall efficiency* which considers all input variables, the *offensive efficiency* which considers only the offensive abilities of the team, and the *defensive efficiency* considering only the defensive abilities.

The efficiency calculations apply to all matches played by each team. However, to look for the presence of the home advantage effect, the researchers also computed the efficiency for matches played at home and matches played away. To summarize, the efficiency calculations are therefore like this: the overall efficiency for all matches, the overall efficiency of a team at home, overall efficiency of a team away, the offensive efficiency for all matches, the offensive efficiency at home, and so on. In total, nine efficiency scores are computed for each of the teams, as shown in Figure 3.

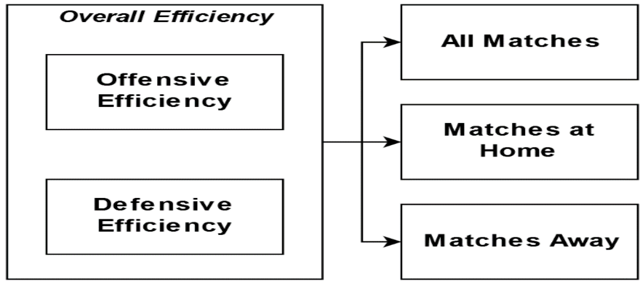

Figure 3. Summary of Efficiency Calculations

The efficiency calculations are done separately for Group A and Group B teams. As discussed previously, the groupings are: Group A: DMUs 1 up to 5; Group B: DMUs 6 up to 10. The researchers used two software-based solvers to computer for the efficiencies: Open Source DEA and DEAP 2.1. Both software supports many types of DEA problems and produce the following solutions. The *DMU score* or *efficiency* is the efficiency score of a DMU relative to its peers. Its values range from 0 to 1. *Projections* or *projected values* are the optimal values of input and/or output variables necessary to reach the efficiency frontier. The *reference set* are DMUs that an inefficient unit can reference in order to improve their own performance. The strength of relationship of an inefficient unit to its reference set is determined by the *lambda* value. These values reflect the relationship of each peer to the DMU. *Slacks* are the potential improvements to the input/output variables that can enable a DMU to reach its *projections*. *Weights* are associated with the input and output variables. There are usually no predefined assigned weights for the input and output variables, but they are computed as part of the optimal solution of DEA. And finally, *efficiency status*. A DMU with an efficiency of 1 is efficient, otherwise it is considered inefficient.

2.6 Correlation

The researchers conducted correlation to understand the linear relationship between two or more variables. The correlation was conducted by considering all the variables as a whole unit without distinguishing whether an input is an offensive or defensive input. The correlation interpretation criteria are as shown in Table 2 (Phanny, 2019).

Table 2. Correlation Coefficient Interpretation Guide

Coefficient Range	Strength Association
$> \pm 0.8$	Very Strong Relationship
$\pm 0.6\ to\ \pm 0.8$	Strong Relationship
$\pm 0.4\ to\ \pm 0.6$	Moderate Relationship
$\pm 0.2\ to\ \pm 0.4$	Weak Relationship
$< \pm 0.2$	Very Weak Relationship

2.7 Regression Analysis

To identify the most impactful input variables that contribute to the generation of outputs, the researchers also conducted regression analysis incorporating the stepwise procedure. The dependent variables are the points and goals, where the regression for each is computed separately. The independent variables are the input variables considered in the study.

The researchers tested the regression for alpha = 0.05 in stepwise regression. The researchers took the R^2 and adjusted R^2 values for Goals and Points as well as the factors computed as significant in stepwise regression model for each alpha. The preferred alpha value is chosen based on the values of the R^2 and adjusted R^2 and the presence of multicollinearity which can be identified by a high variance inflation factor (VIF) value based on the book by Hair et al (2014). The researchers also considered the presence of factors that are not significant i.e. factors in which the p-value of greater than 0.05.

3. Results and Discussion

3.1 Match Results

As discussed in the previous chapter there are two output variables in consideration – the number of Goals and the Points gained. Each team faced each other in a single round match, in which they played at Home or Away for two matches each, a total of 4 matches. Vietnam topped Group A with 8 goals scored overall while Thailand conquered Group B with 15 goals. As for the points earned, for Group A, Vietnam scored the highest with 10 points, followed by Malaysia with 9; for Group B, Thailand came out the top with 10 goals, with the Philippines coming in at second with 8. The top two teams with most points in the group stage advanced to the semi-finals.

For the input variables, the researchers took the mean value of each input variable for all matches, matches at home, and matches away. The offensive input variables display the effective attacking abilities of each team that they used to achieve goals and earn more points. In Group A, Vietnam tied with Malaysia in goal attempts with 15.75. In Group B, Singapore gathered 19 goal attempts, 9 of which are shots on goal during their home games. For the defensive inputs, these inputs are used in football strategy to prevent the opponent from scoring. Laos excels in defense especially with defending corners, producing the highest number for this variable among Group A DMUs. As for Group B, Singapore did not do well in their defensive side which is due to having the lowest average defended corners and goals saved that resulted in highest overall fouls averaging to 19.5 in all matches of Group B.

3.2 Efficiency Results using DEA

The computed efficiencies are discussed in the following order: overall, offensive, and defensive efficiency for all matches, then for matches at home, and finally, for matches away. Table 3 contains the computed overall efficiency scores for both Group A and Group B for all matches. The DMUs have a mean overall efficiency of 91.56% for all matches. Overall efficiency encompasses all variables, inclusive of all offensive and defensive variables.

Table 3. Overall Efficiency, All Matches

Group	DMU No.	Country	Efficiency	Group	DMU No.	Country	Efficiency
A	1	Vietnam	1.000	B	6	Thailand	1.000
	2	Malaysia	1.000		7	Philippines	1.000
	3	Myanmar	1.000		8	Singapore	1.000
	4	Cambodia	0.938		9	Indonesia	0.818
	5	Laos	1.000		10	East Timor	0.400
						Mean	**0.9156**

For Group A, only Cambodia is inefficient relative to its peers in Group A, with an efficiency of 93.8%. This means that Cambodia is not utilizing its resources or inputs effectively to produce output. Cambodia can still improve its efficiency by 6.2% in order to reach the frontier. Despite producing less points and goals than Cambodia, Laos is still efficient since it was able to attain 3 goals at its current level of input (Cambodia only produced 4). Cambodia had significantly greater ball possession, though Laos is better in defending corners, blocked shots, and saves by goalkeeper, and only marginally less fouls (10 vs. 11 for Cambodia). For Group B, Indonesia and East Timor are inefficient, with an efficiency of 81.8% and 40% respectively. This means that Indonesia can still improve its output by 18.2% to reach the frontier, while East Timor can still improve by 60%.

Table 4. Offensive Efficiency, All Matches

Group	DMU No.	Country	Efficiency	Group	DMU No.	Country	Efficiency
A	1	Vietnam	1.000	B	6	Thailand	1.000
	2	Malaysia	1.000		7	Philippines	1.000
	3	Myanmar	1.000		8	Singapore	0.993
	4	Cambodia	0.927		9	Indonesia	0.818
	5	Laos	1.000		10	East Timor	0.400
						Mean	**0.9138**

Table 4 contains the computed offensive efficiency scores for both Group A and Group B for all matches. The offensive efficiency only considers the offensive variables or inputs. The mean offensive efficiency of the DMUs does not differ greatly from the overall efficiencies. For Group A, Cambodia is still the sole inefficient DMU with an efficiency of 92.7% relative to its group. This means that this DMU can still improve by 7.3% in order to reach the frontier. For Group B, the three inefficient DMUs are Singapore, Indonesia and East Timor. The three had an efficiency score of 99.3%, 81.8% and 40% respectively. Compared to Singapore and Indonesia, East Timor had considerably lower efficiency and has the lowest offensive efficiency all teams considered. To reach the frontier, Singapore needs an improvement of 0.7%, Indonesia needs an improvement of 18.2%, while East Timor needs an improvement of 60%. The DMUs had a mean offensive efficiency of 91.38% for all matches.

Table 5 contains the computed defensive efficiency scores for both Group A and Group B in all matches. The defensive efficiency considers only the defensive variables or inputs.

Table 5. Defensive Efficiency, All Matches

Group	DMU No.	Country	Efficiency	Group	DMU No.	Country	Efficiency
A	1	Vietnam	1.000	B	6	Thailand	1.000
	2	Malaysia	0.900		7	Philippines	1.000
	3	Myanmar	0.819		8	Singapore	1.000
	4	Cambodia	0.500		9	Indonesia	0.525
	5	Laos	0.413		10	East Timor	0.285
						Mean	**0.7440**

Compared with the overall and offensive efficiencies before it, the mean defensive efficiency score of all DMUs for all matches is lower, which possibly indicates that the teams have worse performance when it comes to defense. For Group A, only Vietnam is efficient, while Malaysia comes in at second place with 90%, and Laos ranking last with 41.3%. Since Vietnam is the only efficient DMU, it becomes the peer for all the other DMUs in Group A.

For Group B, there are two inefficient DMUs: Indonesia and East Timor. The two had an efficiency score of 52.5% and 28.5% respectively, which is significantly lower than that of the other DMUs in this group. East Timor has the lowest defensive efficiency all teams considered. To reach the frontier, Indonesia needs to improve by 47.5%, while East Timor needs to improve by 71.5%. The mean defensive efficiency of all DMUs for all matches is 74.42%, which is lower than the offensive efficiency by a considerable margin.

Table 6 contains the computed overall efficiency scores for both Group A and Group B DMUs for all matches at home.

Table 6. Overall Efficiency, Matches at Home

Group	DMU No.	Country	Efficiency	Group	DMU No.	Country	Efficiency
A	1	Vietnam	1.000	B	6	Thailand	1.000
	2	Malaysia	1.000		7	Philippines	1.000
	3	Myanmar	1.000		8	Singapore	1.000
	4	Cambodia	1.000		9	Indonesia	1.000
	5	Laos	0.480		10	East Timor	1.000
						Mean	**0.9480**

For Group A, all teams except one is efficient. Laos is the only inefficient DMU, with an efficiency score of 48%. This means that Laos can still improve its efficiency by 52% in order to reach the frontier. For Group B, all DMUs are overall efficient at home, having each scored an efficiency of 100%. Since all the DMUs are efficient in this group, each of them is their own peer. Across both groups, one can say that, with the sole exception of Laos, the DMUs are efficient when playing at home. The mean overall efficiency for matches played at home is 94.8%.

The computed offensive efficiency scores for both Group A and Group B DMUs for matches at home is the exact same as the overall efficiency (see Table 6). As with the overall efficiency, all teams except one is efficient. Laos is the only inefficient DMU, with an efficiency score of 48%. For all DMUs, with the sole exception of Laos, the DMUs are offensively efficient when playing at home. The mean offensive efficiency at home is also 94.8%.

Table 7 contains the computed defensive efficiency scores for both Group A and Group B DMUs for matches at home.

Table 7. Defensive Efficiency, Matches at Home

Group	DMU No.	Country	Efficiency	Group	DMU No.	Country	Efficiency
A	1	Vietnam	1.000	B	6	Thailand	1.000
	2	Malaysia	1.000		7	Philippines	0.833
	3	Myanmar	0.819		8	Singapore	1.000
	4	Cambodia	0.900		9	Indonesia	0.810
	5	Laos	0.244		10	East Timor	0.335
						Mean	**0.7941**

The mean defensive efficiency score at home is lower than that of the overall and offensive efficiencies, with 79.41%. For Group A, only Vietnam and Malaysia are efficient, while Myanmar has an efficiency of 81.9%, Cambodia has an efficiency of 90%, and Laos has the lowest efficiency in the group with only 24.4%. For Group B, there are three inefficient DMUs: Philippines, Indonesia and East Timor. The efficiency scores are 83.3%, 81% and 33.5% respectively. East Timor had the lowest defensive efficiency both for its group and all teams considered.

Table 8 contains the computed overall efficiency scores for both Group A and Group B DMUs for all matches away. For Group A, Cambodia is the only inefficient DMU with an efficiency of 54.8%, which means it can still improve by 45.2% in order to reach the frontier. For Group B, Indonesia and East Timor are still the bottom two teams with an efficiency of 75% and 24.7%, respectively. East Timor scored the lowest efficiency for its group and for all DMUs. The DMUs had a mean overall efficiency of 85.45% for matches away.

Table 8. Overall Efficiency, Matches Away

Group	DMU No.	Country	Efficiency	Group	DMU No.	Country	Efficiency
A	1	Vietnam	1.000	B	6	Thailand	1.000
	2	Malaysia	1.000		7	Philippines	1.000
	3	Myanmar	1.000		8	Singapore	1.000
	4	Cambodia	0.548		9	Indonesia	0.750
	5	Laos	1.000		10	East Timor	0.247
						Mean	**0.8545**

Table 9 contains the computed offensive efficiency scores for both Group A and Group B DMUs for all matches away. The mean offensive efficiency for all matches played away from home is 80.06%, which is slightly lower than the overall efficiency.

Table 9. Offensive Efficiency, Matches Away

Group	DMU No.	Country	Efficiency	Group	DMU No.	Country	Efficiency
A	1	Vietnam	1.000	B	6	Thailand	1.000
	2	Malaysia	1.000		7	Philippines	1.000
	3	Myanmar	1.000		8	Singapore	1.000
	4	Cambodia	0.476		9	Indonesia	0.350
	5	Laos	1.000		10	East Timor	0.180
						Mean	**0.8006**

For Group A, Cambodia is still the only inefficient DMU with an efficiency of 47.6%. For Group B, Indonesia and East Timor are still the worst performers of the group. The efficiency scores of these two inefficient DMUs are significantly far from 100%. Indonesia has an offensive efficiency score of 35% when playing away from home while East Timor scored 18%.

Lastly, Table 10 contains the computed defensive efficiencies of Group A and Group B DMUs for matches played away from home. The mean defensive efficiency of all DMUs for matches away is 74.6%.

Table 10. Defensive Efficiency, Matches Away

Group	DMU No.	Country	Efficiency	Group	DMU No.	Country	Efficiency
A	1	Vietnam	1.000	B	6	Thailand	1.000
	2	Malaysia	1.000		7	Philippines	1.000
	3	Myanmar	0.667		8	Singapore	1.000
	4	Cambodia	0.316		9	Indonesia	0.750
	5	Laos	0.480		10	East Timor	0.247
						Mean	**0.7460**

For Group A, Vietnam and Malaysia are efficient. For the inefficient DMUs, Myanmar has an efficiency of 66.7%, Cambodia has an efficiency 31.6%, and Laos has an efficiency of 48%. For Group B, Indonesia and East Timor are inefficient. Each of them has an efficiency of 75% and 24.7%, respectively. East Timor scored the lowest efficiency for its group and for all DMUs.

3.3 Summary of Efficiency Computations

Table 11 contains the matrix of mean efficiencies for all scenarios considered by the researchers. That is, the overall, offensive, and defensive efficiencies for all matches, matches at home, and finally, matches away.

Table 11. Matrix of Average Efficiencies

Efficiency	All Matches	Home	Away
Overall	0.9156	0.9480	0.8545
Offensive	0.9138	0.9480	0.8006
Defensive	0.7442	0.7941	0.7460

It can be observed that the average overall and offensive efficiencies of the DMUs are higher than the average defensive efficiency. This observation is true considering all matches and for matches played at home and away. This might indicate the need for the DMUs to improve the utilization of their defensive abilities or inputs. It can also be observed that on average, the DMUs are more efficient at home than away. This observation points toward the presence of a *home advantage effect*. This hypothesis is supported by similar results from the study by Villa and Lozano [6]. The reason for this is that football teams generally perform better in their home field due to factors like familiarity with the football pitch and the psychological effect of crowd response to the team's performance. At home, offense is considerably greater than defense, which necessitates the need to improve the teams' defensive abilities at home. On the other hand, when playing away, the discrepancy between offense and defense is lower, but teams can benefit in being more offensive. The study of Bosca et al (2009) which studied Italian and Spanish football leagues produced similar results, in which they recommended the Spanish leagues to be offensive at home and away and for Italian leagues to be offensive at home and defensive away.

3.4 Correlation

For all matches, many input variables share a very strong relationship. For instance, ball possession and pass accuracy have a very strong relationship because in a football game, a team that has a high percentage of ball possession usually controls the tempo of the game, the key strategy being precision in terms of passes between players. The more precise the passes, the higher the ball possession which is crucial in a build-up play to score a goal. This result is consistent with the study of Zambom-Ferraresi et al (2018), which placed a higher importance in accuracy of passes. Attacks and dangerous attacks also have a very strong relationship; technically dangerous attacks are considered a type of attack that happened in the critical areas of the pitch. The more attacking attempts made by the offensive team the higher the chance they could create dangerous attacks given that the build-up play is situated inside the penalty box. Meanwhile, shots on goal has a very strong relationship to goal attempts. The reason behind this is that the aim of every offensive strategy to continue to increase those amounts, and for that strategy to be effective every attempt should yield large amount of shots towards the goal. Another very strong relationship is between ball possession and defended corners. One of the critical chances of the offensive team to score a goal is through creative ball possession, thus defending to ball coming from the build-up play in the corner is a must for the defending team.

For home matches, three variable pairs yield a very strong relationship among input variables namely: ball possession to pass accuracy, ball possession to dangerous attacks and dangerous attacks to defended corners. a great build-up play starts from a good handling of the ball. Excellent ball possession begins with great passing skill, the less error the better. Once the build-up play is excellent an offensive team can effortlessly enter the danger zone (penalty box) and create an attack in attempt to score a goal. Every offensive team will make the most out of their chance of converting a dangerous attack especially during a tie breaker. Frequently, the team who has the ball is set to execute a set piece based on their game strategy. Thus, the defensive team will also do their part in preventing the ball from entering the net.

Unlike the statistical correlation results for all matches and matches played at home, which has more moderate input variable relationship, the results of correlation for matches played away has a greater number of strong correlated relationship to every input variable. Even in an away match, the relationship of ball possession to pass accuracy, shots on goal to goal attempts and dangerous attacks to attacks all are still very strong. This means that those attacking traits are still being maintained by all the DMUs during their away matches. The relationship of ball possession to pass accuracy in away matches is lower compared to home matches, but still has a very strong relationship. One possible reason for this is the atmosphere of away pitch, which is different from home, which is why players could commit errors in passing which affect the execution of their build-up play. Many teams in Europe practices counter-attacking strategy during away game, this shows that fast paced creation of attacks is needed. The tendency is a variety of attacks and dangerous attacks will occur which will as a result of the increase in goal attempts as well as shots on goal.

3.5 Regression

The regression model of points generated by football teams in their team standings (using stepwise method) for all matches consist of two independent variables comprised of pass accuracy and shots on goal. These independent variables account for an 82.11% of the variability of the regression model as indicated by the coefficient of determination, R^2 or 76.99% as indicated by the adjusted R^2. The regression equation is: ***Points$_{ALL}$ = -19.17 + 25.27 Pass Accuracy + 1.286 Shots on Goal.*** As shown in the regression equation, the increase in pass accuracy and shots on goal result to an increase in points. The regression implies that in a football match, the more precise passes of the ball, more chances to score will be created by the forward players which will eventually increase the number of shots on goal. A team could create more shots on goal when precise passing is executed flawlessly among players, given that the forward players perform a greater number of forward runs towards the goal.

The regression model for the goals generated by football teams for all matches also consist of two independent variables: shot accuracy and corner kicks. These independent variables account for an 80.08% of the variability of the regression model as indicated by the coefficient of determination, R^2 or 74.38% as indicated by the adjusted R^2. The regression equation is: ***Goals$_{ALL}$ = -7.40 + 25.36 Shot Accuracy + 0.924 Corner Kick.*** As shown in the regression equation, the increase in shot accuracy and corner kicks result to an increase in goals. In a football match, when a player tries as many scoring attempts as he can, attempts at goal either from the danger area or the corner of the box increase the chances of him scoring a goal. However, for a team to win the game they also need to improve on practicing set pieces from the corner kicks in order to utilize every chance they created not just in the build-up play but also the other chances like corner kick to secure a goal.

The regression model of points generated by football teams in their team standings (using stepwise method) for all matches at home consist of two independent variables comprised of pass accuracy and shots on goal. These independent variables account for an 87.23% of the variability of the regression model as indicated by the coefficient of determination, R^2 or 83.58% as indicated by the adjusted R^2. The regression equation is: ***Points$_{HOME}$ = -9.41 + 13.87 Pass Accuracy + 0.569 Shots on Goal.*** The regression model of points for matches at home has the exact same significant variables as the model for all matches. At home, a team could create more shots on goal when precise passing is executed flawlessly among players, given that the forward players perform a greater number of forward runs towards the goal.

The regression model for the goals at home consist of only one independent variable: shots on goal. This independent variable account for an 73.61% of the variability of the regression model as indicated by the coefficient of determination, R^2 or 70.31% as indicated by the adjusted R^2. The regression equation is: ***Goals$_{HOME}$ = 0.508 + 0.671 Shots on Goal.*** Usually, the team that plays in their home ground has more attempts, off or on goal. More shots taken by the players could boost the team's confidence in creating a threat that would ensure that the shot will land in the back of the net, scoring a goal.

The regression model of points generated by football teams for all matches away consist of three independent variables comprised of ball possession, blocked shots, and goal saves. These independent variables account for an 84.38% of the variability of the regression model as indicated by the coefficient of determination, R^2 or 76.57% as indicated by the adjusted R^2. The regression equation is: ***Points$_{AWAY}$ = -4.37 + 9.26 Ball Possession + 1.054 Blocked - 0.574 Goal Saves.*** As shown in the regression equation, the increase in ball possession and blocked shots result to an increase in points, while a decrease in goal saves result to an increase in points. A common strategy in an away game is maintaining ball possession of the teams, blocking shots, and preventing goal attempts as much as possible.

Finally, the regression model for the goals generated by football teams for all matches away consist of only one variable: fouls. The regression equation is: ***Goals$_{AWAY}$ = 7.32 - 0.387 Fouls.*** As shown in the regression equation, the decrease in fouls result to an increase in goals. In matches away from home, the abilities of both teams are uncertain in terms of execution of playing strategy thus, outcome is unpredictable. For the away team to secure victory they usually tighten their defensive abilities, like blocking shots, standing tackle etc. Thus, unintentionally they can commit fouls that aims to prevent the home team from scoring a goal.

4. Conclusion

This paper evaluated the technical efficiency of Southeast Asian football teams that competed in the AFF Suzuki Cup 2018 using Data Envelopment Analysis (DEA), under the CCR assumption (constant returns to scale) and with an output-orientated approach. The researchers chose the DEA approach to calculate the efficiency frontier – that is, the maximum achievable performance of each team, to identify which teams are offensive and defensively efficient in their strategy with respect to the chosen input variables (ball possession, goal attempts, etc.) and output variables (points and goals produced). With DEA, the researchers were able to benchmark the football teams in order to discover the best-performing teams and identify gaps in output for the inefficient teams, in line with similar football DEA studies (Villa and Lozano, 2016), (Bosca, 2009), (Rubem, 2015). Finally, the researchers were able to identify the relationship between input and output variables and how they interact with and affect each other, to find out which variables contribute significantly to a team's success.

References

Alegi, P. C., Joy, B., Weil, E., Rollin, J., and Giulianotti, R. C. Football-Soccer. Retrieved from https://www.britannica.com/sports/football-soccer.

Anderson, M., Wolfson, S., Neave, N., Moss, M., Perspectives on the home advantage: A comparison of football players, fans and referees. Psychology Sport and Exercise, 13(3), 311–6, 2012, doi.org/10.1016/j.psychsport.2011.11.012

ASEAN Football Federation, AFF Suzuki Cup 2018 Tournament Regulations, p. 1 – 32, 2018.

Bialik, C. The People Tracking Every Touch, Pass And Tackle in the World Cup, *FiveThirtyEight*. 2014, Available from: https://fivethirtyeigbht.com/features/the-people-tracking-every-touch-pass-and-tackle-in-the-world-cup/

Bosca, J. E., Liern V., Martinez A., and Sala, R. Increasing offensive or defensive efficiency? An analysis of Italian and Spanish football. *Omega*, 37(1), 63 – 78, 2009.

Charnes, A., Cooper, W. W., Rhodes, E., Measuring the efficiency of decision making units. *European Journal of Operational Research*. 2(6), 429–44, 1978.

Chitnis A., Vaidya, O., Performance assessment of tennis players : Application of DEA. *Procedia Social and Behavioral Science,* 133:74–83, 2014., doi.org/10.1016/j.sbspro.2014.04.171

Cooper, W. W., Seiford, L. M., Zhu, J., *Handbook on data envelopment analysis*, Springer, U.S., 2004. doi: 10.1007/b105307

Football (Soccer)/The Basics, *Wikibooks, open books for an open world*, Retrieved from https://en,wikibooks.org/wiki/Football_(Soccer)/The Basics.

Hair, J. F. J., Black, W. C., Babin, B. J., Anderson, R.E., Multivariate Data Analysis. 7th Edition. *Pearson Education Limited*, Pearson Education Limited, 2014.

Kumar, S. and Gulati, R, An Examination of Technical, Pure Technical, and Scale Efficiencies in Indian Public Sector Banks using Data Envelopment Analysis, *Eurasian Journal of Business and Economics*, 1 (2), 33 – 69, 2008.

Lewis, H. F., Lock, K. A., Sexton, T. R., Organizational capability, efficiency, and effectiveness in Major League Baseball: 1901 - 2002, *European Journal of Operational Research*, 197 (2), 731 – 740, 2009. doi.org/10.1016/j.ejor.2008.07.002.

Lovett, G., Nielsen Sports World Football Report 2018, 1–35, 2018, Retrieved from https://nielsensports.com/wp-content/uploads/2014/12/Nielsen_World-Football-2018-6.11.18.pdf

Nevill, A., Balmer, N., Williams, M, Crowd influence on decisions in association football. *Lancet,* 353(9162), 1416, 1999.

Phanny I. Guideline for interpreting correlation coefficient [Internet]. [cited 2019 Oct 21]. Available from: https://www.slideshare.net/phannithrupp/guideline-for-interpreting-correlation-coefficient

Rogge, N., Puyenbroek, T. V., Reeth, D. V., Performance evaluation of Tour de France cycling teams using Data Envelopment Analysis, *International Journal of Sport Finance*, 8 (3), 2012.

Rubem, A. P. S., Brandão, L. C., Multiple criteria data envelopment analysis - An application to UEFA EURO 2012. Procedia Compuert Science, 2015;55(Itqm):186–95. Available from: http://doi.org/10.1016/j.procs.2015.07.031

Ruiz, L., Sirvent, I., Cooper, W. W., Selecting non-zero weights to evaluate effectiveness of basketball players with DEA, *European Journal of Operational Research*, 195 (2), 563 – 574, 2009.

Villa, G., Lozano, S., Assessing the scoring efficiency of a football match, European Journal of Operational Research, 255 (2), 559 – 69, 2016, https://doi.org/10.1016/j.ejor.2016.05.024.

Vincova, K., Using DEA Models to Measure Efficiency, BIATEC, Volume XIII, 8/2005.

Wu H., Hsiao C., Chiu S., Measuring Pitchers' Performance Using Data Envelopment Analysis with Advanced Statistics. Contemporary Management Research, 11(4), 351–84, 2015, 11.10.7903/cmr.14157.

Zambom-Ferraresi F., Rios, V. Lera-Lopez, F. Determinants of sport performance in European football: What can we learn from the dat?, Decis Support Syst, 114 (August), 18 – 28, 2018.

Biographies

Rene D. Estember is currently a Professor in the School of Industrial Engineering and Engineering Management at the Mapua University in Manila City, Philippines. He earned his B.S. in Management and Industrial Engineering from Mapua Institute of Technology in 1979, Master in Business Administration from Ateneo de Manila University in 1994, Master of Science in Industrial Engineering from the University of the Philippines in 2008. He is also a Professional Industrial Engineer certified by the Philippine Institute of Industrial Engineers in 2008 and an ASEAN Engineer. He has 17 years of work experiences in the industry from 1979 up to 1996 while teaching part-time from 1992 up to 2000 in various schools. He is also providing consultancy services and conducting technical trainings. His research interests include human factors and ergonomics, manufacturing, risk management and optimization. He has published conference papers indexed in Scopus. He is an active member of the Operations Research Society of the Philippines (ORSP), Philippine Institute of Industrial Engineers (PIIE) and the Mapua Association of Management Service and Industrial Engineers (MAMSIE).

Shean Michael L. Reyes finished his B.S. in Industrial Engineering degree from Mapua University.

Omar A. Solaiman finished his B.S. in Industrial Engineering degree from Mapua University.