

# Assessing Offensive/Defensive Strategies in a Football Match Using DEA

G. Villa<sup>1</sup> and S. Lozano<sup>1</sup>

<sup>1</sup>*University of Seville*

**Gabriel Villa** is an associate professor in the Department of Industrial Management at the University of Seville. He holds a PhD from the University of Seville, and his primary research interest is in data envelopment analysis. He has published papers related to sports in high impact journals, including soccer (*European Journal of Operational Research*), American football (*Annals of Operations Research*), and Olympic Games and basketball (*Journal of the Operational Research Society*).

**Sebastián Lozano** is professor of quantitative methods in the Department of Industrial Management at the University of Seville. He holds a PhD from the University of Seville and a Master of Engineering in industrial management from Katholieke Universiteit Leuven. His research interests include efficiency assessment and complex networks analysis. He has published papers applying these techniques to sports, including applications to Olympic Games, soccer, American football, basketball, and F1 auto racing.

## Abstract

Football teams play following two strategies: defensive and offensive. The length of time in which the team plays under each of these two modes will depend on who the rival team is and the technical play developed. Once the match is over, an analysis of the team's performance is necessary to determine which aspects of the offensive/defensive strategies used during the game failed and which succeeded. Although there are some studies in the literature that assess the offensive and defensive efficiencies of football teams in a season, none of them deal with the performance in a single match. In this paper, a multiple modes of functioning (MMF) data envelopment analysis (DEA) model is proposed. This methodology determines the percentage of time in which the team should have played in defensive and offensive modes in order to maximize those efficiencies. The procedure has been applied to the Spanish First Division league in the 2014/2015 season.

Keywords: performance assessment, offensive/defensive strategies, DEA, multiple modes of functioning

<http://doi.org/10.32731/IJSF.143.082019.01>

## Introduction

Data envelopment analysis (DEA) is a well-known nonparametric methodology aimed at benchmarking similar operating units. These units are assumed to consume inputs and produce outputs. DEA allows for determining whether the observed level of outputs could have been achieved with a lower level of inputs or, alternatively, whether a larger amount of outputs could have been produced with the observed input consumption. DEA has been widely used in multiple applications related to sports (Li et al., 2015; Wu et al., 2010; Lozano et al., 2003; Moreno & Lozano, 2015; etc.), including football (e.g., Villa & Lozano, 2016).

In this paper, however, we are not interested in assessing the performance of teams during a season, as most DEA studies do. We focus our attention on a specific match, taking into account the team data as well as data from the opposing team. Analyzing how a team plays is important to understand its success or its failure in a match. Basically, two strategies can be deployed during a match: the offensive game (when the team has the possession of the ball) and the defensive game (when it does not). Thus, we aim to assess the efficiency of the offensive/defensive strategies of each team. Moreover, we propose a novel DEA approach recently used in Lozano and Villa (2016) and Lozano et al. (2017) for systems with multiple modes of functioning (MMF). The working of that type of system results from the interplay of different subprocesses. The key feature of MMF DEA, one that distinguishes it from network DEA, for example, is the fact that the subprocesses do not run parallel; instead, the operation of the system alternates between the different subprocesses, which thus run on a time-sharing basis.

In this paper, the offensive and defensive strategies of the game played by a team are considered its two distinct modes of functioning (MF). In each of them, the team uses certain inputs to produce outputs. Thus, for example, goals scored is

the output of the offensive strategy while goals against (a variable that the team is interested in reducing as much as possible) is the input of the defensive MF. An important variable in MMF DEA is the time allocation between the different MF. That corresponds, in this application, to the fraction of time the team is in possession of the ball. Not only can the proposed approach compute efficient targets for each strategy (and corresponding efficiency scores), but it can also determine the optimal time allocation. Thus, this analysis provides a more detailed and rich analysis than that of the conventional, single-process, black box (BB) DEA approach with which it is also compared in this paper.

The structure of the paper is as follows. Section 2 carries out a literature review of DEA applied to football teams. In Section 3, we formulate and explain the proposed MMF DEA models. In Section 4, the proposed approach is illustrated using the different games played by the 20 football teams of the Spanish First Division in the 2014/2015 season. Finally, Section 5 summarizes and concludes.

## Literature Review

Many studies point to a direct relationship between the success of a team and the strategic mix used during a certain number of games played. Some empirical studies focused on the success and failure of football teams are as follows. Hughes et al. (1988), who examined patterns of play for successful (semifinals) and unsuccessful (eliminated) teams in the 1986 World Cup finals, concluded that successful teams played significantly more touches of the ball per possession than unsuccessful teams. Hughes and Franks (2005) studied the relationship between number of passes and goals in the 1990 and 1994 World Cup finals. Tenga et al. (2010) deduced that counterattacks were more effective than elaborate attacks when playing against an imbalanced defense, using a sample of 163 matches played in the Norwegian professional football league. Lago and Martín (2007) used a sample of 170 matches in the First Division of the Spanish League to study the relationship between winning, drawing, or losing and the percentage of the possession of the ball. Taylor et al. (2008) focused their study on 40 games of a professional English team, concluding that in winning it performed more interceptions, clearances, and aerial challenges and fewer crosses, passes, and dribbles, while in losing it made more crosses, dribbles, and passes and fewer clearances and interceptions.

As regards to the DEA methodology used in this paper, several papers have studied the performance of football teams in order to assess their efficiency. Haas et al. (2003a) was the first study in measuring the efficiency of football teams using DEA. Since then, many other studies have addressed this issue from different perspectives. Espitia-Escuer and García-Cebrián (2004, 2006, 2008, 2010) studied Spanish football teams using different approaches (CCR-I, CCR-O, BCC-I, and BCC-O models). Guzmán and Morrow (2007) used the directors' remuneration of English Premier League teams in order to explain the sport's results measured through points won and turnovers. Sala et al. (2009) used Windows Analysis to measure the efficiencies evaluated from the 2000–01 to 2007–08 seasons. Recently, Zambom-Ferraresi et al. (2017) analyzed the technical, pure technical, and scale efficiencies of Europe's top 32 international football competitions at club level through 10 sports seasons (2003–04 to 2013–14). They used as economic output the coefficients applied by the UEFA from UCL revenue distribution.

To group all these DEA studies, we have followed the approach in García-Sánchez (2007) that considered three categories, depending on how the analysis has been carried out empirically: considering economic variables, using the available precise statistical data, and taking into account the emotional aspects. Table 1 shows the variables used by the different DEA references divided into these three categories. Thus, "operating cost," "team payroll," "coach salary," and "season total revenues" are examples of economic variables most often used by the authors; "number of spectators per match" is the most used in the emotional variables category; and finally, the empirical category includes variables such as "number of points," "goals scored," "shots at goal," and "possession." From this classification, the approach proposed in this paper mainly uses empirical variables plus the team and opposing team budgets as additional economic-type variables.

Note that García-Sánchez (2007) was the first to explicitly measure offensive and defensive efficiencies in order to determine the social effectiveness of a team during a season. Later, Boscá et al. (2009) analyzed the offensive and defensive efficiencies in the Italian and Spanish leagues from seasons 2000/2001 to 2002/2003. To do so, they used the average values of goals scored and conceded, shots on/at goal, attacking plays, center plays/in area, attacks in area, and possession of the ball collected from each season.

**Table 1. Review of the Variables Used in the Literature on Performance Assessment of Football Teams Using DEA**

	Variable	References
ECONOMIC	operating cost	Barros and García-del-Barrio (2011); Guzmán and Morrow (2007); Sala et al. (2009)
	team value	Villa and Lozano (2016)
	total assets	Barros and García-del-Barrio (2011)
	team payroll	Barros and García-del-Barrio (2011); Barros and Leach (2006); Haas (2003b); Haas et al. (2004)
	attendance receipts	Barros and García-del-Barrio (2011)
	other receipts	Barros and García-del-Barrio (2011)
	net assets	Barros and Leach (2006)
	expenditures	Barros and Leach (2006)
	staff costs	Guzmán and Morrow (2007)
	directors' remuneration	Guzmán and Morrow (2007)
	total wages and salaries (excl. coach)	Haas (2003a)
	UEFA's coefficient revenue distribution	Zambom-Ferraresi et al. (2017)
	coach salary	Haas (2003a); Haas (2003b); Haas et al. (2004)
	season total revenues	Haas (2003a); Haas (2003b); Haas et al. (2004)
EMOTIONAL	stadium facilities	Barros and Leach (2006)
	stadium capacity	García-Sánchez (2007)
	population of the teams' home town	García-Sánchez (2007); Haas (2003a)
	number of spectators per match	Barros and Leach (2006); García-Sánchez (2007); González-Gómez and Picazo-Tadeo (2010); Haas (2003b); Haas et al. (2004); Roboredo et al. (2015)
EMPIRICAL	number of players	Barros and Leach (2006); Espitia-Escuer and García-Cebrián (2004); Espitia-Escuer and García-Cebrián (2006); Espitia-Escuer and García-Cebrián (2008); Espitia-Escuer and García-Cebrián (2010); González-Gómez and Picazo-Tadeo (2010); Picazo-Tadeo and González-Gómez (2009)
	points	Barros and Leach (2006); Espitia-Escuer and García-Cebrián (2004); Espitia-Escuer and García-Cebrián (2006); Espitia-Escuer and García-Cebrián (2008); González-Gómez and Picazo-Tadeo (2010); Guzmán and Morrow (2007); Haas (2003a); Haas (2003b); Haas et al. (2004); Picazo-Tadeo and González-Gómez (2009); Roboredo et al. (2015)
	goals scored	Boscá et al. (2009); Espitia-Escuer and García-Cebrián (2004); Espitia-Escuer and García-Cebrián (2008); García-Sánchez (2007); Villa and Lozano (2016)
	shots at goal	Boscá et al. (2009); Espitia-Escuer and García-Cebrián (2006); Espitia-Escuer and García-Cebrián (2008); García-Sánchez (2007); Villa and Lozano (2016)
	attacking moves	Espitia-Escuer and García-Cebrián (2004); Espitia-Escuer and García-Cebrián (2006); Espitia-Escuer and García-Cebrián (2008); Espitia-Escuer and García-Cebrián (2010); García-Sánchez (2007)
	center plays/in area	Boscá et al. (2009)
	possession	Boscá et al. (2009); Espitia-Escuer and García-Cebrián (2004); Espitia-Escuer and García-Cebrián (2006); Espitia-Escuer and García-Cebrián (2008); Espitia-Escuer and García-Cebrián (2010); Villa and Lozano (2016); Zambom-Ferraresi et al. (2017)
	shots and headers	Espitia-Escuer and García-Cebrián (2004)
	goals attempts	Espitia-Escuer and García-Cebrián (2008); Espitia-Escuer and García-Cebrián (2010); Zambom-Ferraresi et al. (2017)
	passes	Zambom-Ferraresi et al. (2017)
passes to the penalty area	García-Sánchez (2007)	

Continued

**Table 1. Review of the Variables Used in the Literature on Performance Assessment of Football Teams Using DEA**  
(continued from page 133)

EMPIRICAL	ball recovery	García-Sánchez (2007); Villa and Lozano (2016); Zambom-Ferraresi et al. (2017)
	goalkeeper's actions	García-Sánchez (2007); Villa and Lozano (2016)
	number of seasons played in the First Division	González-Gómez and Picazo-Tadeo (2010); Picazo-Tadeo and González-Gómez (2009); Roboredo et al. (2015)
	trophies won in competitions	González-Gómez and Picazo-Tadeo (2010)
	number of matches played in European competitions	Espitia-Escuer and García-Cebrián (2010); González-Gómez and Picazo-Tadeo (2010); Picazo-Tadeo and González-Gómez (2009)
	number of matches played in the King's Cup	González-Gómez and Picazo-Tadeo (2010); Picazo-Tadeo and González-Gómez (2009)
	attacks/in area	Boscá et al. (2009)
	center plays/in area	Boscá et al. (2009)
	turnover	Barros and Leach (2006); Guzmán and Morrow (2007); Villa and Lozano (2016)
	corners and penalties	Villa and Lozano (2016)
	inverse of goals received	Boscá et al. (2009); García-Sánchez (2007)
	position in the final league table	García-Sánchez (2007); Roboredo et al. (2015)

In addition, most previous DEA applications studied the performance of teams in a season. However, we believe that once a given match is over, an analysis of the game played is essential in order to identify the strengths and weaknesses of one's own team, with the aim of improving for the next match. To the best of our knowledge, only Villa and Lozano (2016) have analyzed the efficiency of football teams in a single match. Specifically, they used a network DEA model with two parallel processes: the home team and the away team. However, they did not study the offensive/defensive strategies used by each team.

In this paper, we will consider that a football team can be defined as a system with two MFs: offense and defense. When the offense mode is used, the team cannot operate in the defense mode, and vice versa. The possession of the ball determines the percentage of the total time that the team functions in offense mode during the match. Also, the inputs and outputs involved when the team attacks are different from the inputs and outputs to be considered when the team defends. To deal with this situation, we propose using the MMF DEA approach used by Lozano and Villa (2016). This approach allows measuring the offensive and defensive efficiencies of a team in a given match and computing targets for the number of goals scored and against. Furthermore, the proposed approach also determines the optimal time allocation between the two strategies.

### Proposed DEA Approach

Recently, Lozano and Villa (2016) proposed a DEA approach to model situations in which the system can operate in one of the MMF. In this section, we present the application of the MMF DEA approach to the case of a football team playing a given match. Figure 1 shows a football team considered from a BB perspective as well as distinguishing its two MFs. Note that the inputs and outputs of the BB approach are the sum of those of the offense and defense modes.

Inputs in the offense mode are nondiscretionary since the amount of each one is not only determined by the team but also by the rival. Similarly, the outputs of the defense mode must be considered as nondiscretionary for the same reasons. In the offense MF, the team tries to convert as many shots at goals, corners, penalties, etc. into goals scored, while in the defense mode, it tries to prevent the shots at goals against, corners against, penalties against, etc. to be converted into goals against. Note that the inputs and outputs of the whole system are the sum of the inputs and outputs of its two MFs and that, except for the budget variables, the set of inputs and outputs of the two MFs are disjointed. The budget of the team has been considered as a nondiscretionary input in both MFs, since it contributes positively to their performance but is

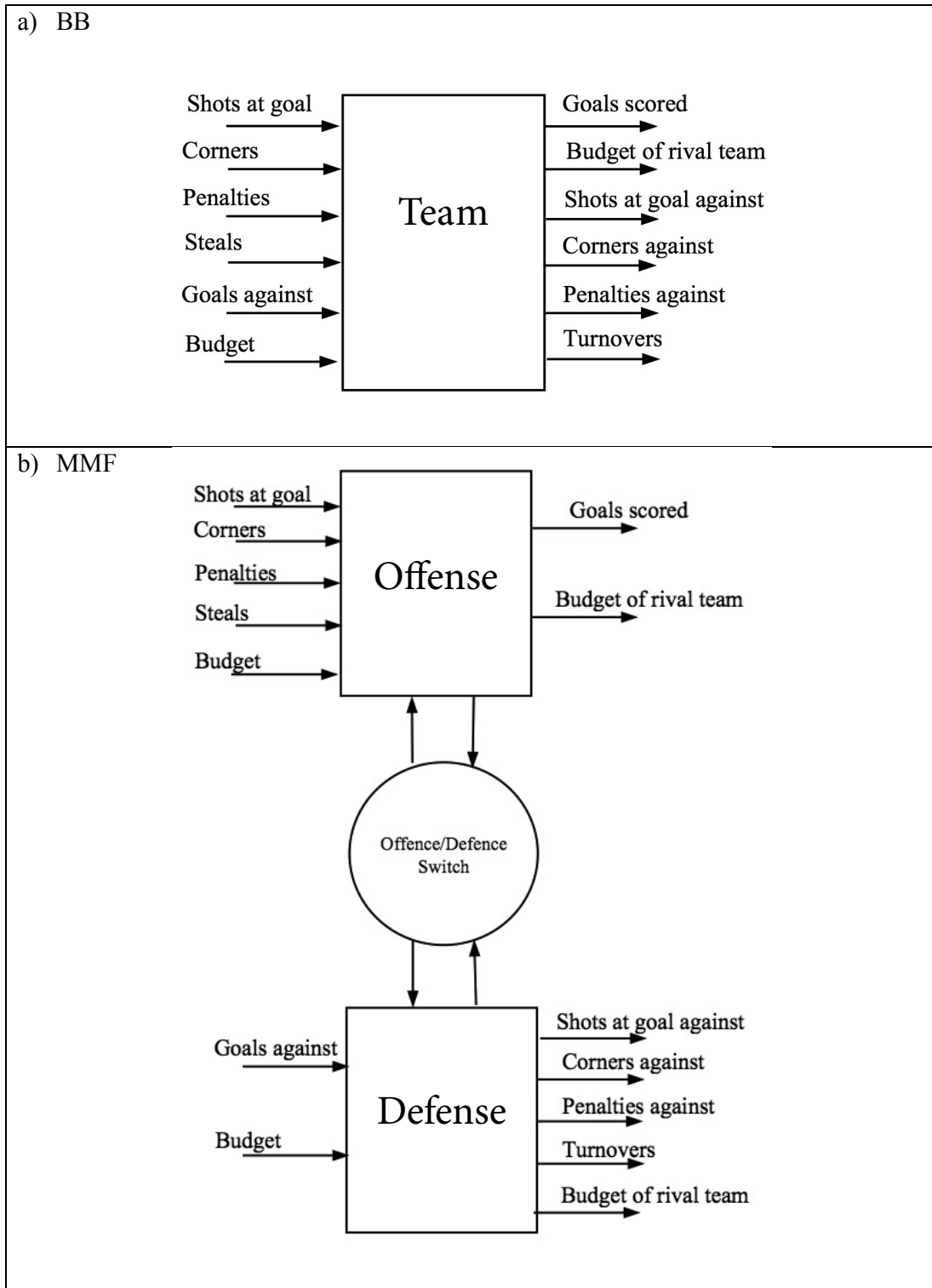
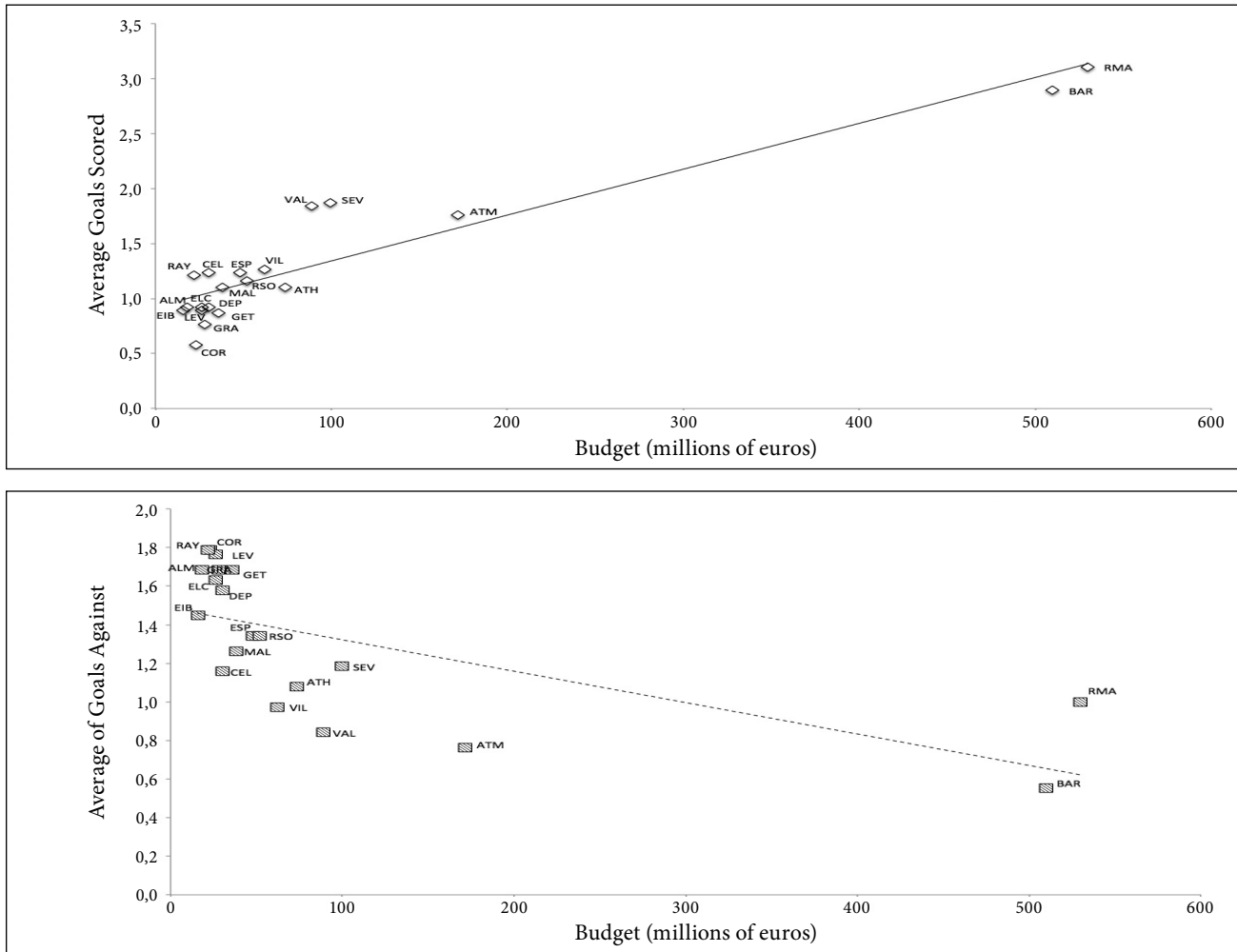


Figure 1. BB and MMF perspectives for performance analysis of football team in a match

considered fixed during a season. Similarly, the budget of the rival is considered as a nondiscretionary output in both MFs. These budget variables have been considered because there seems to be a correlation between the budget of a team, the average number of goals scored, and the average number of goals against during the season, with wealthier teams scoring more and receiving fewer goals against than more modest teams. Thus, Figure 2 shows those variables for the 20 teams of the Spanish First Division for the season 2014–2015.



**Figure 2. Average number of goals scored and of goals against versus budget**

Before formulating the BB and MMF DEA models, let us introduce the required notation:

$I^{off}$  set of non-budget offensive inputs: shots at goal, corners, penalties, and steals.

$O^{def}$  set of non-budget defensive outputs: shots at goal against, corners against, penalties against, and turnovers.

$x_{ij}^{off}$  amount of input  $i \in I^{off}$  consumed in offense mode by DMU  $j$ .

$y_{kj}^{def}$  amount of output  $k \in O^{def}$  produced in defense mode by DMU  $j$ .

$pos_j$  ball possession (i.e., fraction of time in offense mode) for DMU  $j$ .

$gs_j$  number of goals scored by DMU  $j$ .

$ga_j$  number of goals against for DMU  $j$ .

$b_j$  budget of DMU  $j$ .

$br_j$  budget of the rival team of DMU  $j$ .

**BB DEA Model**

Conventional DEA (e.g., Cooper et al., 2000) does not distinguish MF, so it therefore cannot take into account that the team can only play in either offense or defense mode. Thus, the BB DEA model represents the traditional perspective of

assessing the efficiency by considering the game played by a team in a football match as a BB. The key feature of the BB DEA model is that, as is customary in conventional DEA, a single set of intensity variables is used to project the DMU. Thus:

$(\lambda_1, \lambda_2, \dots, \lambda_n)$  intensity variables

sgs slack for goals scored by DMU 0

sga slack for goals against DMU 0

$$\begin{aligned} & \text{BB} \\ & \text{Max sgs} + \text{sga} \end{aligned} \tag{1}$$

s.t.

$$\sum_j \lambda_j x_{ij}^{\text{off}} \leq x_{i0}^{\text{off}} \quad \forall i \in I^{\text{off}} \tag{2}$$

$$\sum_j \lambda_j b_j \leq b_0 \tag{3}$$

$$\sum_j \lambda_j ga_j = ga_0 - \text{sga} \tag{4}$$

$$\sum_j \lambda_j gs_j = gs_0 + \text{sgs} \tag{5}$$

$$\sum_j \lambda_j y_{kj}^{\text{def}} \geq y_{k0}^{\text{def}} \quad \forall k \in O^{\text{def}} \tag{6}$$

$$\sum_j \lambda_j br_j \geq br_0 \tag{7}$$

$$\sum_j \lambda_j = 1 \tag{8}$$

$$\lambda_j \geq 0 \quad \forall j \text{ sga, sgs integer} \tag{9}$$

The objective function of the above model corresponds to a non-oriented, variable returns to scale (VRS), additive DEA model. Note that the objective function tries to increase goals scored and reduce goals against at the same time. Improving on any of these dimensions contributes to the overall objective of winning the match. The nondiscretionary inputs and outputs do not use slack variables and, hence, those inputs and outputs are not included in the objective function. Finally, since the goals scored and goals against targets should be integer variables, the corresponding slacks are integers (Lozano & Villa, 2006; Kuosmanen & Kazemi-Matin, 2009; Kazemi-Matin & Kuosmanen, 2009).

#### MMF1 DEA Model

The proposed MMF DEA approach distinguishes two alternative MFs when a team is playing: the offensive and defensive game. Each MF has its own set of intensity variables, which allows for computing its target operation point within its own mode-specific technology. Such MF target points are computed as linear combinations of the corresponding MF of the observed DMUs. This is done first by computing the input consumption rates and output production rates of each DMU and then by using the corresponding intensity variables to combine them. The input consumption rate represents

the input consumption per unit of time. Since the unit of time considered is the duration of the whole game, the input consumption rate of a DMU corresponds to the input consumption if the given MF has been played by that DMU all the time. Similarly, the output production rate of a DMU corresponds to the amount of output that the DMU had obtained if it had played that MF all the time. For each MF, the intensity variables represent the fraction of time that the DMU 0 being assessed should replicate the corresponding MF operation point of the observed DMUs—the same as happens in a conventional DEA, in the optimal solution, only efficient DMUs will have nonzero intensity variables. In other words, the MF targets can only result from linearly combining DMUs that are efficient and thus define the facet of the MF efficient frontier on which the MF target lies.

Let:

$(\lambda_1^{off}, \lambda_2^{off}, \dots, \lambda_n^{off})$  intensity variables for offense mode

$(\lambda_1^{def}, \lambda_2^{def}, \dots, \lambda_n^{def})$  intensity variables for defense mode

MMF1

$$\begin{aligned} & \text{Max } sgs + sga & (10) \\ & \text{s.t.} \end{aligned}$$

$$\sum_j \lambda_j^{off} \frac{x_{ij}^{off}}{pos_j} \leq x_{i0}^{off} \quad \forall i \in I^{off} \tag{11}$$

$$\sum_j \lambda_j^{off} \frac{b_j}{pos_j} \leq b_0 \tag{12}$$

$$\sum_j \lambda_j^{def} \frac{ga_j}{(1-pos_j)} = ga_0 - sga \tag{13}$$

$$\sum_j \lambda_j^{def} \frac{b_j}{(1-pos_j)} \leq b_0 \tag{14}$$

$$\sum_j \lambda_j^{off} \frac{gs_j}{pos_j} = gs_0 + sgs \tag{15}$$

$$\sum_j \lambda_j^{off} \frac{br_j}{pos_j} \geq br_0 \tag{16}$$

$$\sum_j \lambda_j^{def} \frac{y_{kj}^{def}}{(1-pos_j)} \geq y_{k0}^{def} \quad \forall k \in O^{def} \tag{17}$$

$$\sum_j \lambda_j^{def} \frac{br_j}{(1-pos_j)} \geq br_0 \tag{18}$$



$$\sum_j \lambda_j^{\text{off}} = \text{pos}_0 \quad (19)$$

$$\sum_j \lambda_j^{\text{def}} = 1 - \text{pos}_0 \quad (20)$$

$$\lambda_j^{\text{off}}, \lambda_j^{\text{def}} \geq 0 \quad \forall j \text{ sga, sgs integer} \quad (21)$$

Note that, although MMF1 has the same objective function as BB, the input and output constraints (11)–(18) are different from those of the BB model in two respects. One is that two different sets of intensity variables are used, one for each MF. The other is that, since the system can only operate in one of the two MFs at any time and the fraction of time each MF is used is given by the observed possession percentage, the MF operation points result from operating a given fraction of time (given by the intensity variables) as the different observed DMUs. The MF target inputs and outputs are computed by aggregating the inputs consumed and outputs produced during those fractions of time. The overall sum of all the intensity variables is unity, meaning that the sum of the length of time that the team plays in offense and defense modes is equal to the total match duration.

As indicated above, a key feature of the MMF1 DEA model is that the time allocation, i.e., the time that each MF is used, is fixed to the observed value, as per equations (19) and (20). This constraint will be relaxed in the alternative MMF DEA model presented below.

#### **MMF2 DEA Model**

In the MMF2 DEA model, as in the previous model, the intensity variables determine, for each MF, the fraction of time that the team should play as each of the observed DMUs. The difference with respect to MMF1 is that, in MMF2, the model is free to compute the most efficient time allocation, i.e., the fraction of time for both MFs that leads to the largest potential improvement in terms of increase in goals scored and decrease in goals against.

Let:

$\alpha^{\text{off}}$  fraction of time that DMU 0 should be in offense mode

$\alpha^{\text{def}}$  fraction of time that DMU 0 should be in defense mode

The MMF2 DEA model coincides with the MMF1 model by just replacing constraints (19) and (20) with the following:

$$\sum_j \lambda_j^{\text{off}} = \alpha^{\text{off}} \quad (22)$$

$$\sum_j \lambda_j^{\text{def}} = \alpha^{\text{def}} \quad (23)$$

$$\alpha^{\text{off}} + \alpha^{\text{def}} = 1 \quad (24)$$

$$\alpha^{\text{off}}, \alpha^{\text{def}} \geq 0 \quad (25)$$

Note that since any feasible solution in MMF1 is also feasible in MMF2 (but not the opposite), the optimal solution of MMF2 involves a larger (or at least equal) value of the objective function, which is the sum of the increase in goals scored and reduction in goals against.

#### **Illustration of Proposed Approach**

In this section, the application of the proposed approach to the performances of the teams of the Spanish First Division in each of the matches played during the 2014/2015 season is presented. Tables 2 and 3 show the average value of the inputs and outputs of each MF observed for each team in the different matches played during that season.

**Table 2. Average of the Inputs and Outputs of Football Teams Operating in Offense Mode**

OFFENSE MODE	INPUTS					OUTPUTS		
TEAM	# SHOTS AT GOAL	# CORNERS	# PENALTIES	# STEALS	BUDGET (10 <sup>6</sup> €)	GOALS SCORED	BUDGET OF RIVAL TEAM (10 <sup>6</sup> €)	% POSSESSION
ALMERIA	3.18	5.24	0.16	40.68	18.3	0.92	50.29	46.04
ATHLETIC	3.89	5.21	0.13	47.24	73.532	1.11	48.84	52.09
ATLETICO	4.97	4.89	0.16	47.08	171.7	1.76	46.25	49.06
BARCELONA	7.24	6.16	0.18	50.82	509.6	2.89	37.36	68.84
CELTA	4.37	6.24	0.16	51.34	30.3	1.24	49.98	57.11
CORDOBA	3.24	4.47	0.08	41.87	23.071	0.58	50.17	36.86
DEPORTIVO	3.50	4.61	0.18	42.79	30.08	0.92	49.98	46.52
EIBAR	3.29	4.24	0.05	43.18	15.8	0.89	50.36	40.92
ELCHE	3.39	4.37	0.18	42.00	26.4	0.92	50.08	32.17
ESPANYOL	3.63	4.53	0.08	45.50	48.2	1.24	49.50	49.66
GETAFE	3.84	4.79	0.05	45.55	36	0.87	49.83	49.06
GRANADA	3.34	5.29	0.16	38.13	28	0.76	50.04	50.12
LEVANTE	2.68	4.68	0.05	42.16	26.253	0.89	50.08	49.91
MALAGA	4.42	5.16	0.16	46.18	38	1.11	49.77	50.78
R MADRID	7.47	6.24	0.32	48.84	529.5	3.11	36.84	50.45
R SOCIEDAD	3.55	5.32	0.08	46.34	51.872	1.16	49.41	50.49
RAYO	4.53	5.16	0.08	48.37	21.967	1.21	50.19	50.87
SEVILLA	4.95	5.21	0.24	45.89	99.7	1.87	48.15	50.00
VALENCIA	4.32	4.58	0.18	46.84	89	1.84	48.43	50.39
VILLARREAL	4.68	5.68	0.05	45.74	62.099	1.26	49.14	49.57

**Table 3. Average of the Inputs and Outputs of Football Teams Operating in Defense Mode**

DEFENSE MODE	INPUTS		OUTPUTS					
TEAM	# GOALS AGAINST	BUDGET (10 <sup>6</sup> €)	# SHOTS AT GOAL AGAINST	# CORNERS AGAINST	# PENALTIES AGAINST	# TURNOVERS	BUDGET OF RIVAL TEAM (10 <sup>6</sup> €)	1-POSSESSION (%)
ALMERIA	1.68	18.3	5.71	6.82	0.05	66.00	50.29	53.96
ATHLETIC	1.08	73.532	3.76	4.84	0.08	74.74	48.84	47.91
ATLETICO	0.76	171.7	2.58	4.26	0.08	64.66	46.25	50.94
BARCELONA	0.55	509.6	2.45	3.63	0.11	65.76	37.36	31.16
CELTA	1.16	30.3	3.92	3.79	0.16	72.95	49.98	42.89
CORDOBA	1.79	23.071	5.13	5.68	0.18	67.82	50.17	63.14
DEPORTIVO	1.58	30.08	4.84	4.82	0.16	72.92	49.98	53.48
EIBAR	1.45	15.8	5.16	6.00	0.13	69.08	50.36	59.08
ELCHE	1.63	26.4	4.71	5.71	0.24	62.74	50.08	67.83
ESPANYOL	1.34	48.2	4.39	5.03	0.03	72.11	49.50	50.34
GETAFE	1.68	36	4.53	5.21	0.21	70.50	49.83	50.94
GRANADA	1.68	28	4.24	4.50	0.11	63.71	50.04	49.88
LEVANTE	1.76	26.253	4.89	5.32	0.21	63.34	50.08	50.09
MALAGA	1.26	38	4.03	4.58	0.11	70.21	49.77	49.22
R MADRID	1.00	529.5	3.58	4.92	0.13	68.82	36.84	49.55
R SOCIEDAD	1.34	51.872	4.55	6.26	0.16	64.32	49.41	49.51
RAYO	1.79	21.967	4.84	4.97	0.13	73.76	50.19	49.13
SEVILLA	1.18	99.7	3.95	4.66	0.21	66.26	48.15	50.00
VALENCIA	0.84	89	3.37	5.61	0.21	66.32	48.43	49.61
VILLARREAL	0.97	62.099	3.87	5.45	0.05	67.97	49.14	50.43

The three DEA models presented in Section 2 were solved for the 760 DMUs (corresponding to each of the 20 teams playing 38 matches during the season). Figure 3 shows the absolute frequency counts for the pairwise comparisons of the optimal value of the objective function computed by the three models. For points above the diagonal line, the optimal objective function of the approach corresponding to the Y-axis is larger than the optimal objective function value of the approach corresponding to the X-axis. The opposite occurs for points below the diagonal. Note that a larger optimal objective function value means more inefficiency identified by the corresponding model. As expected, the optimal objective function values for MMF1 are always less than (or equal to) those of MMF2. As regards BB, its optimal objective function value is larger than that of MMF1 a few times and larger than that of MMF2 in just two cases. It follows that MMF2 has more discriminant power than MMF1 and the latter more discriminant power than BB.

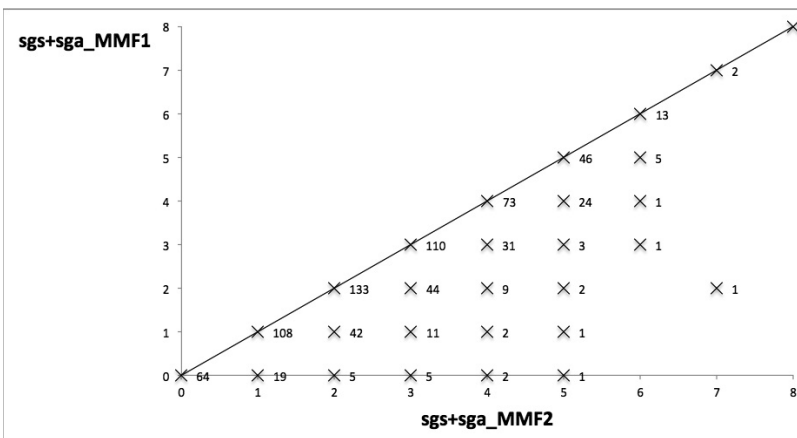
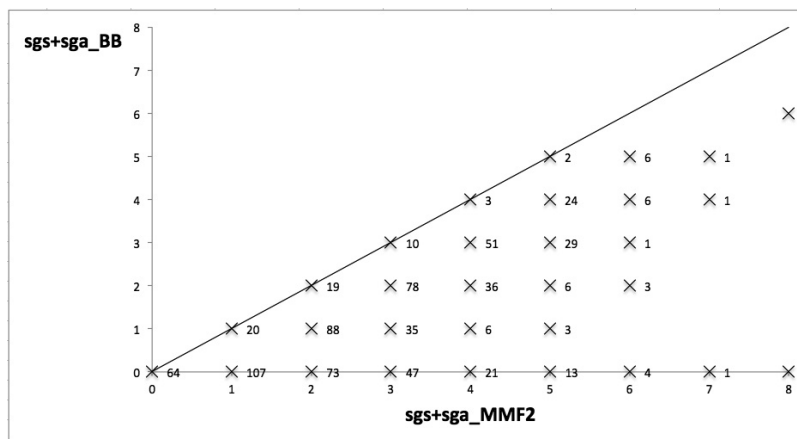
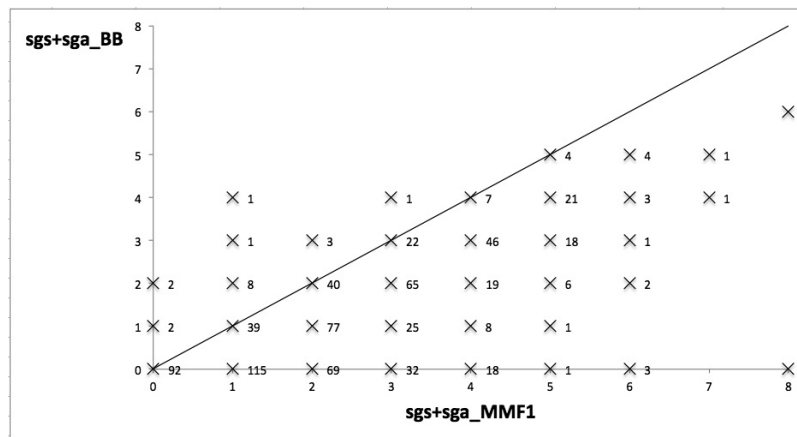


Figure 3. Sum of goals-scored and goals-against slacks computed by BB and MMF models

The objective function of the proposed models provides scores that measure inefficiency and are not normalized. If normalized efficiency scores are desired, these can be computed for any of the three approaches (BB, MMF1 and MMF2) as

$$\theta_0 = \frac{1}{1 + sgs^* + sga^*} \tag{26}$$

$$\theta_{0,off} = \frac{0.5}{0.5 + sgs^*} \tag{27}$$

$$\theta_{0,def} = \frac{0.5}{0.5 + sga^*} \tag{28}$$

It is easy to check that the overall system efficiency  $\bar{\theta}$  is the harmonic mean of offensive ( $\theta_{0,off}$ ) and defensive ( $\theta_{0,def}$ ) efficiencies, i.e.,

$$\frac{1}{\theta_{0,off}} + \frac{1}{\theta_{0,def}} = \frac{1}{0.5} \left( \frac{1 + sgs^* + sga^*}{1} \right) = \frac{2}{\theta_0} \Leftrightarrow \theta_0 = \frac{2}{\frac{1}{\theta_{0,off}} + \frac{1}{\theta_{0,def}}} \tag{29}$$

Table 4 shows, for each team, the average value of the overall, offensive, and defensive efficiencies (27) –(29), labelled as  $\bar{\theta}$ ,  $\bar{\theta}_{off}$ , and  $\bar{\theta}_{def}$ , respectively, and the average value of the targets of goals scored ( $gs_0 + sgs^*$ ) and goals against ( $gs_0 + sga^*$ ) computed by each of the three approaches considered. It can be seen that, generally,  $\bar{\theta}^{BB} > \bar{\theta}^{MMF1} > \bar{\theta}^{MMF2}$ ,  $\bar{\theta}_{off}^{BB} > \bar{\theta}_{off}^{MMF1} > \bar{\theta}_{off}^{MMF2}$ , and  $\bar{\theta}_{def}^{BB} > \bar{\theta}_{def}^{MMF1} > \bar{\theta}_{def}^{MMF2}$ , which suggests, as mentioned above, that MMF2 is the approach with the greatest discriminant power. Note also that, for most teams, the defensive efficiency is higher than the offensive efficiency. This is more frequent for MMF2, where it occurs for 17 out of the 20 teams.

**Table 4. Average Overall, Offense, and Defense Efficiency and Goals-Scored and Goals-Against Targets Computed by BB and MMF Models**

TEAM	$\bar{\theta}$			$\bar{\theta}_{off}$			$\bar{\theta}_{def}$			gs+sgs*			ga-sga*		
	EM	MMF1	MMF2	EM	MMF1	MMF2	EM	MMF1	MMF2	EM	MMF1	MMF2	EM	MMF1	MMF2
ALMERIA	0.644	0.451	0.398	0.793	0.595	0.517	0.685	0.539	0.485	1.32	2.42	1.97	0.92	0.21	0.45
ATHLETIC	0.507	0.434	0.391	0.510	0.449	0.414	0.698	0.629	0.560	2.39	2.82	2.55	0.5	0.55	0.11
ATLETICO	0.501	0.404	0.378	0.476	0.396	0.348	0.760	0.653	0.653	3.13	2.39	3.61	0.32	0.45	0.03
BARCELONA	0.559	0.418	0.361	0.500	0.338	0.295	0.846	0.832	0.753	4.37	2.95	5.26	0.24	0.55	0.05
CELTA	0.532	0.411	0.344	0.554	0.423	0.357	0.806	0.611	0.507	2.42	2.53	2.79	0.74	0.29	0.18
CORDOBA	0.536	0.413	0.336	0.593	0.522	0.378	0.675	0.483	0.437	1.45	2.66	2.13	0.97	0.45	0.32
DEPORTIVO	0.555	0.416	0.344	0.695	0.543	0.421	0.611	0.525	0.472	1.55	2.82	2.32	0.74	0.21	0.16
EIBAR	0.800	0.546	0.498	0.869	0.697	0.573	0.844	0.632	0.613	1.16	2.47	1.63	1.11	0.18	0.63
ELCHE	0.644	0.408	0.382	0.748	0.558	0.507	0.715	0.481	0.458	1.47	2.61	2.16	0.92	0.37	0.24
ESPANYOL	0.527	0.375	0.350	0.590	0.475	0.460	0.648	0.503	0.493	2.13	2.58	2.55	0.58	0.26	0
GETAFE	0.492	0.292	0.269	0.524	0.364	0.335	0.632	0.398	0.354	1.97	2.97	2.53	0.89	0.53	0.18
GRANADA	0.576	0.375	0.327	0.745	0.525	0.425	0.641	0.480	0.454	1.34	2.79	2.03	0.97	0.37	0.16
LEVANTE	0.573	0.400	0.379	0.656	0.544	0.506	0.692	0.457	0.415	1.47	2.76	1.92	0.97	0.37	0.29
MALAGA	0.491	0.382	0.333	0.592	0.409	0.378	0.635	0.519	0.460	2.16	2.47	2.79	0.47	0.45	0.08
R. MADRID	0.481	0.301	0.290	0.435	0.301	0.296	0.754	0.586	0.576	4.68	2.47	5.26	0.47	0.26	0.05
R. SOCIEDAD	0.465	0.389	0.329	0.524	0.513	0.407	0.649	0.487	0.453	2.26	2.79	2.53	0.55	0.34	0.11
RAYO	0.626	0.529	0.329	0.742	0.549	0.390	0.710	0.716	0.507	1.68	2.63	2.63	0.95	0.26	0.34
SEVILLA	0.546	0.365	0.316	0.550	0.423	0.398	0.770	0.621	0.530	2.87	2.47	3.34	0.71	0.45	0.03
VALENCIA	0.566	0.477	0.450	0.591	0.496	0.465	0.725	0.648	0.630	2.74	2.55	3.18	0.34	0.26	0.16
VILLARREAL	0.412	0.309	0.292	0.389	0.354	0.324	0.728	0.612	0.590	2.82	2.58	3.19	0.34	0.34	0.03

With respect to the average target values, all teams present a higher value of goals scored than goals against for the three approaches. The range of values of the average target values of goals scored is larger (e.g., between 1.63 and 5.26 for MMF2) than the corresponding range for the target goals against (e.g., between 0.0 and 0.63 for MMF2). Note also how the targets of the top teams (such as Barcelona and Real Madrid) are more demanding, especially for goals scored. Finally, through its ability to optimize the time allocation, MMF2 may sometimes be rather optimistic, especially as regards the target goals against, which is often close to zero.

Following the suggestion of one of the reviewers, we have tested whether the obtained results were robust with respect to deleting the matches played by the two top teams, Real Madrid and Barcelona. Tables 5 and 6 show, respectively, the average and Pearson correlation of the efficiency scores of the different DEA models for the complete dataset and for the dataset without Real Madrid and Barcelona. As can be seen from those tables, the results obtained in both cases are very similar. This must be due to the fact that the budget variable, which is the one that sets these two teams apart from the rest, is considered nondiscretionary and, therefore, the proposed approach benchmarks the small teams against similar small teams.

**Table 5. Average Efficiency Scores with and without Real Madrid and Barcelona**

		EFF	EFF_OFF	EFF_DEF
BB DEA model	Data without RM and BAR	0.343	0.390	0.500
	Complete Data	0.343	0.389	0.504
MMF1 DEA model	Data without RM and BAR	0.414	0.489	0.564
	Complete Data	0.395	0.458	0.555
MMF2 DEA model	Data without RM and BAR	0.635	0.708	0.739
	Complete Data	0.601	0.668	0.721

**Table 6. Pearson Correlation Coefficient of Efficiency Scores with and without Real Madrid and Barcelona**

	EFF	EFF_OFF	EFF_DEF
BB DEA model	0.929	0.870	0.828
MMF1 DEA model	0.946	0.928	0.979
MMF2 DEA model	0.981	0.983	0.985

Figure 4 shows the value of the optimal time allocation  $\alpha^{off}$  (i.e., the fraction of time in a match that the team should have played in an offensive mode) computed by MMF2 versus the actual, observed possession of the ball for each of the 760 DMUs. The points with  $\alpha^{off} > pos_0$  (above the diagonal line) correspond to cases in which teams should play in the offensive mode longer, while for points below the diagonal line (which occur more often), the opposite happens (i.e., the team should play for longer in a defensive mode). The symbol distinguishes whether the DMU won, drew, or lost the corresponding match. Note that the points above and farthest from the diagonal tend to be wins. Those points correspond to the common situation in which a winning team plays in defensive mode (trying to maintain its advantage and leaving the initiative to the rival), contrary to the more offensive game suggested by the MMF2 model. Analogously, some points below and farthest from the diagonal line correspond to teams that lost or drew the match. Those points correspond to situations in which a team is losing or drawing and, trying to improve that score, the MMF2 model suggests a more offensive game than the one actually played.

Figure 4 also shows that, according to the adage “the best defense is a good offense,” it may seem counterproductive to increase the share of the defense MF. In defensive mode, a team cannot make goals (and so cannot increase their output) while they are at risk of receiving goals (increasing their inputs). Recall that the MMF DEA objective function involves two slacks, one for goals scored and the other for goals against. Increasing the possession of the ball aims at increasing the

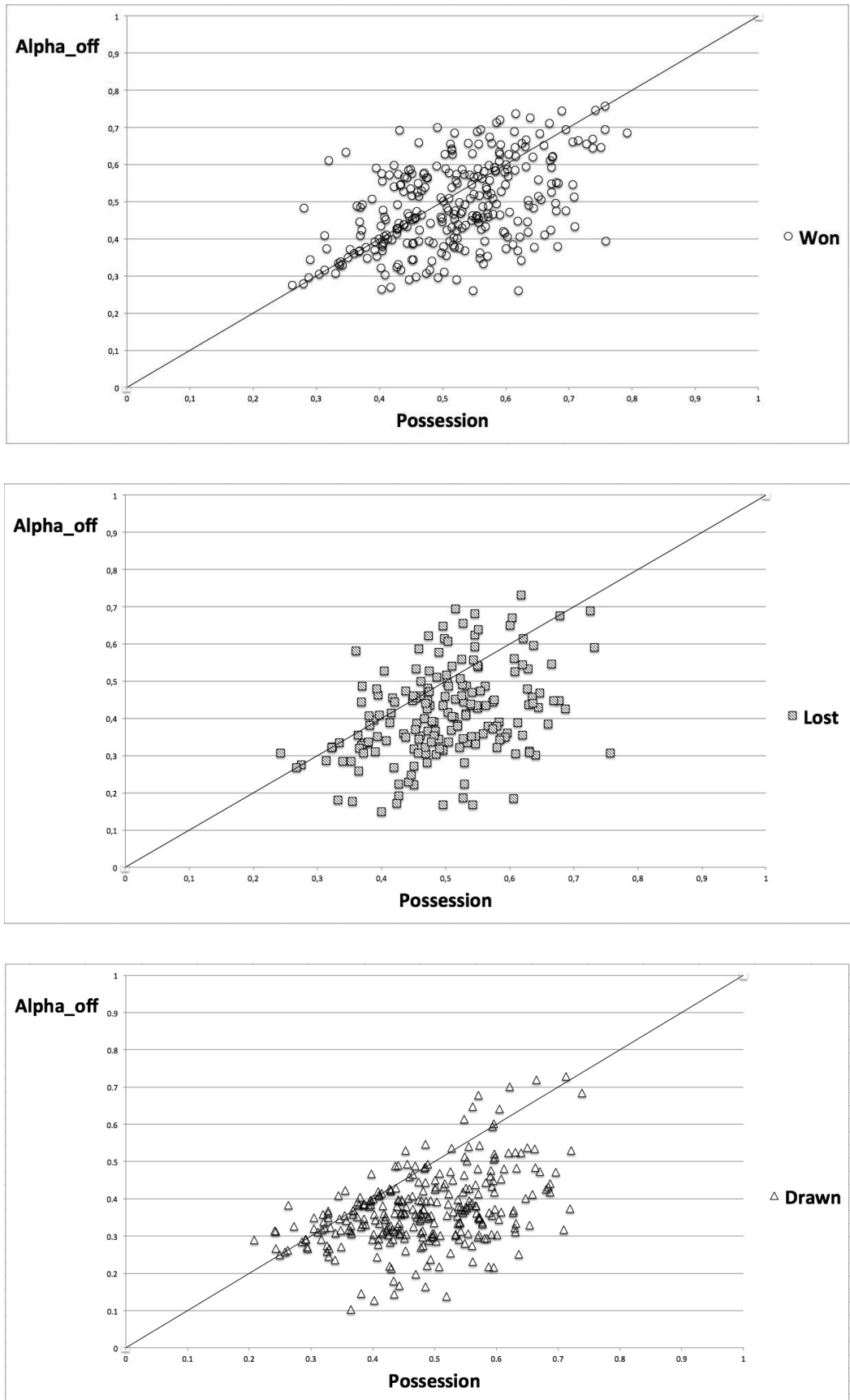


Figure 4. Observed possession versus optimal  $\alpha^{off}$  for the 760 DMUs grouped by win, lose, or draw

goals scored, which would positively impact the objective function. However, by playing an efficient defense game, a team may reduce its goals against, especially if the observed DMU had many goals against, perhaps because it played too much offense, disregarding its defense. In that case, more defense can be beneficial for the objective function.

## Summary and Conclusions

In this paper, a new MMF DEA approach is used to assess the offensive/defensive strategies of every single match played by a team. To do so, the team has been modeled as a system that works under two MFs: the offense mode and the defense mode. Depending on whether the time allocated to each MF is respected or optimized, two different models (labelled MMF1 and MMF2, respectively) are proposed. The proposed approach can compute MF efficiency scores as well as MF targets and optimal MF time allocation. For comparison, the conventional BB DEA model has also been considered. It is important to note that the main contribution of this article to the existing football DEA literature is this new perspective that seeks to determine the optimal fractions of time for both strategies of the game (defense and offense) as well as target improvements in goals scored and reductions in the goals against, which, respectively, measures the inefficiencies in the effectiveness of these two strategies.

The proposed approach has been applied to the teams in the Spanish First Division for the matches they played during the 2014/2015 season. The results obtained by the proposed MMF DEA models have been compared with the BB DEA model, and it has been confirmed that the proposed approach has more discriminant power than conventional DEA and makes better use of the available information on how the real system works. Overall and mode-specific efficiency scores have been computed for each team in each match, together with target goals scored and goals against. The optimal offensive/defensive time allocation has also been computed and compared with the observed possession of the team.

As regards limitations of the study, the model proposed in this article may not be applicable in certain situations where the best strategy to be played may be influenced not just by the variables considered but also by other external factors (e.g., the score of matches played by third teams or by the score of the previous match in knock-out rounds).

For possible further topics of research, it would be very interesting to consider additional variables such as the ranking of the DMUs at the time the match is played to analyze if the team ranked below in the table tends to play in a defense mode most of the time (and the opposite for the team ranked above in the table). Another interesting research line would be using a real-time data collection system. Thus, it would be possible to incorporate these models into a useful tool capable of monitoring the game and gathering data from the match (such as the remaining duration of the game and the current score) so as to decide to change the initial strategy if necessary in order to improve the team performance. On the other hand, this procedure can be applied to other sports where the game is based on defensive and offensive tactics, such as, for example, basketball.

## References

- Barros, C. P., & García-del-Barrio, P. (2011). Productivity drivers and market dynamics in the Spanish first division football league. *Journal of Productivity Analysis*, 35, 5–13.
- Barros, C. P., & Leach, S. (2006). Performance evaluation of the English Premier Football League with data envelopment analysis. *Applied Economics*, 38, 1449–1458.
- Boscá, J. E., Liern, V., Martínez, A., & Sala, R. (2009). Increasing offensive or defensive efficiency? An analysis of Italian and Spanish football. *Omega*, 37, 63–78.
- Cooper, W. W., Seiford, L. M., & Tone, K. (2000). *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Norwell, MA: Kluwer Academic Publishers.
- Espitia-Escuer, M., & García-Cebrián, L. I. (2004). Measuring the efficiency of Spanish first-division soccer teams. *Journal of Sport Economics*, 5(4), 329–346.
- Espitia-Escuer, M., & García-Cebrián, L. I. (2006). Performance in sports teams: Results and potential in the professional soccer league in Spain. *Management Decision*, 44(8), 1020–1030.
- Espitia-Escuer, M., & García-Cebrián, L. I. (2008). Measuring the productivity of Spanish first division soccer teams. *European Sport Management Quarterly*, 8(3), 229–246.
- Espitia-Escuer, M., & García-Cebrián, L. I. (2010). Measurement of the efficiency of football teams in the champions league. *Managerial and Decision Economics*, 31, 373–386.
- García-Sánchez, I. M. (2007). Efficiency and effectiveness of Spanish football teams: A three-stage-DEA approach. *Central European Journal of Operations Research*, 15, 21–45.
- González-Gómez, F., & Picazo-Tadeo, A. J. (2010). Can we be satisfied with our football team? Evidence from Spanish professional football. *Journal of Sport Economics*, 11(4), 418–442.

- Guzmán, I., & Morrow, S. (2007). Measuring efficiency and productivity in professional football teams: Evidence from the English Premier League. *Central European Journal of Operations Research*, 15, 309–328.
- Haas, D. J. (2003a). Productive efficiency of english football teams: A data envelopment analysis approach. *Managerial and Decision Economics*, 24, 403–410.
- Haas, D. J. (2003b). Technical efficiency in the major league soccer. *Journal of Sports Economics*, 4, 203–215.
- Haas, D., Kocher, M. G., & Sutter, M. (2004). Measuring efficiency of German football teams by data envelopment analysis. *Central European Journal of Operations Research*, 12(3), 251–268.
- Hughes, M. D., & Franks, I. (2005). Analysis of passing sequences, shots and goals in soccer. *Journal of Sports Sciences*, 23(5), 509–512.
- Hughes, M. D., Robertson, K., & Nicholson, A. (1988). An analysis of the 1984 world cup of association football. In T. Reilly, A. Lees, K. Davids, & W. Murphy (Eds.), *Science and Football* (pp. 363–367). London: E & FN Spon.
- Kazemi-Matin, R., & Kuosmanen, T. (2009). Theory of integer-valued data envelopment analysis under alternative returns to scale axioms. *Omega*, 37(5), 988–995.
- Kuosmanen, T., & Kazemi-Matin, R. (2009). Theory of integer-valued data envelopment analysis. *European Journal of Operational Research*, 192, 658–667.
- Lago, C., & Martín, R. (2007). Determinants of possession of the ball in soccer. *Journal of Sports Sciences*, 25(9), 969–974.
- Li, Y., Lei, X., Dai, Q., & Liang, L. (2015). Performance evaluation of participating nations at the 2012 London Summer Olympics by a two-stage data envelopment analysis. *European Journal of Operational Research*, 243, 964–973.
- Lozano, S., & Villa, G. (2006). Data envelopment analysis of integer-valued inputs and outputs. *Computer and Operations Research*, 33(10), 3004–3014.
- Lozano, S., & Villa, G. (2016). *Data envelopment analysis of systems with multiple modes of functioning* (Working Paper No. 76076). Retrieved from <https://mpira.ub.uni-muenchen.de/76076/>
- Lozano, S., Villa, G., & Eguía, I. (2017). Data envelopment analysis with multiple modes of functioning: Application to reconfigurable manufacturing systems. *International Journal of Production Research*, 55(24), 7566–7583.
- Lozano, S., Villa, G., Guerrero, F., & Cortés, P. (2003). Measuring the performance of nations at the Summer Olympics using data envelopment analysis. *Journal of the Operational Research Society*, 53(5), 501–511.
- Moreno, P., & Lozano, S. (2015). Estimation of productivity change of NBA teams from 2006–07 to 2012–13 seasons. *International Journal of Sport Finance*, 10(3), 217–241.
- Roboredo, M. C., Aizemberg, L., & Meza, L. A. (2015). The DEA game cross efficiency model applied to the Brazilian football championship. *Procedia Computer Science*, 55, 758–763.
- Sala, R., Liern, V., Martínez, A., & Boscá, J. E. (2009). Analysis and evolution of efficiency in the Spanish Soccer League (2000/01–2007/08). *Journal of Quantitative Analysis in Sports*, 5(1), 1–24.
- Taylor, J., Mellalieu, S., James, N., & Shearer, D. (2008). The influence of match location, quality of opposition, and match status on technical performance in professional association football. *Journal of Sports Sciences*, 26(9), 885–895.
- Tenga, A., Holme, I., Ronglan, L., & Bahr, R. (2010). Effect of playing tactics on goal scoring in Norwegian professional soccer. *Journal of Sports Sciences*, 28(3), 237–244.
- Villa, G., & Lozano, S. (2016). Assessing the scoring efficiency of a football match. *European Journal of Operational Research*, 255, 559–569.
- Wu, J., Zhou, Z., & Liang, L. (2010). Measuring the performance of nations at Beijing Summer Olympics using integer-valued DEA model. *Journal of Sport Economics*, 11(5), 549–566.
- Zambom-Ferraresi, F., García-Cebrián, L. I., Lera-López, F., & Iráizoz, B. (2017). Performance evaluation in the UEFA champions league. *Journal of Sport Economics*, 18(5), 448–470.

## Acknowledgments

The research of the authors has been supported by the Spanish Ministry of Science, Innovation and Universities under the Projects PGC2018-095786-B-I00 and MINECO-18-DPI2017-85343-P.