Received: 15 November 2021 | Revised: 21 March 2022 | Accepted: 25 June 2022

DOI: 10.1111/sms.14206

ORIGINAL ARTICLE

WILEY

Deciphering the offensive process in women's elite football: A multivariate study

Rubén Maneiro[1] | **Iyán Iván-Baragaño**[2] | **José L. Losada**[3] | **Antonio Ardá**[2]

[1]Department of Science of Physical Activity and Sport, Pontifical University of Salamanca, Salamanca, Spain

[2]Department of Physical and Sport Education, University of A Coruña, A Coruña, Spain

[3]Department of Social Psychology and Quantitative Psychology, University of Barcelona, Barcelona, Spain

Correspondence
Rubén Maneiro, Department of Science of Physical Activity and Sport, Pontifical University of Salamanca, Salamanca, Spain.
Email: rubenmaneirodios@gmail.com

Over the last few years, there has been considerable increase in scientific knowledge about women's football. However, the tactical and tactical-strategic aspects have not yet been sufficiently covered in scientific literature. Therefore, this work proposed the following aims: To describe how the offensive phase is produced in women's football, to identify the significant statistical criteria that may be modulating success in attack, and finally to propose different predictive success models, with the ultimate aim of passing this knowledge on to the applied field. The observational methodology was used, one of the most appropriate methodologies for the analysis of motor behaviors in sport. The units of analysis collected and analyzed were 6063 attacks carried out during the FIFA Women's World Cup Canada 2015 and France 2019. The available results demonstrate that, on the one hand, offensive team actions are ineffective (almost 70% finish unsuccessfully), but criteria such as the start form of the attack, zone of ball possession, partial match result, or ball possession time are statistically significant criteria that modulate attack success (goal, shot or pass into the area). Lastly, the multivariate results allow us to propose a theoretical model, passing the probability of success from 31% in the absence of a model, to a theoretical auction probability of 52.6%, based on fast attacks with the intervention of few players, and with possession zone in the opposite field. These results could be directly transferred to the practical field where trainers and technical bodies can put this information into practice in training sessions or matches.

KEYWORDS
decision trees, offensive phase, performance analysis, possession outcome, Women's football

1 | INTRODUCTION

Women's football has grown considerably in the last few years. According to data from the latest FIFA Research Report 2019, nearly 14 million girls and women play organized football throughout the world, of which 4 million are federated and linked to official organizations and tournaments. According to a recent study,[1] in countries where women's football is actively promoted, women's national football teams do perform better.

Scientific knowledge of women's football is currently expanding.[1,2] Trainers and technical bodies are beginning to have access to the first scientific studies relating to this football that allow them to increase and strengthen their players' performance. However, this growth has not been homogenous in the different fields that make up sporting

performance, but has grown disproportionally in certain areas,[2,3] focussing above all on the physical and physiological aspects of the game.[3,4]

In terms of technical and tactical aspects, the available literature is still extremely scarce. The scientific transfer to football teams and sporting organizations in the form of publications is still in need of a boost, both in quantity and quality.[5] One of the most important aspects of football is to score more goals than your opponent, and to achieve this, it is necessary to establish successful offensive behavior. In men's football, the scientific literature states that 60%–70% of goals come from attacking mechanisms which start with ball possession.[6] However, in women's football, there is hardly any work that has covered the subject of attack or the offensive phase.

What little work there is has focussed on specific aspects such as accurate passes, where to regain possession, or the influence of match status in the tactical behavior of the female players. In this sense, the work of Scanlan et al[7] found that in order to achieve shot on goal situations, possession should be regained during dynamic play (possession following stealing or interception), and not via an interruption in play; on the contrary, other studies[8,9] analyzed the technical performance of the players in different FIFA Women's World Cups, concluding that winning teams made a higher number of accurate passes, more shots on goal, and had better ball recuperation patterns than losing teams. The study, carried out by Maneiro et al,[10] analyzed how match status influenced the teams' tactical behavior, concluding that successful teams had more ball possession when they were winning, while the losing teams had more ball possession in their own half, and when they were losing.

Regarding performance and offensive effectiveness, Iván-Baragaño et al.[11] analyzed the offensive phase in the FIFA Women's World Cup 2019 and proposed a predictive offensive success model based on ball possession in the attacking midfield zone with the tactical intention of progressing toward the rival goal, results which were also supported by other similar studies[12,13]; on the contrary, the study carried out by Ibáñez et al.[14] concluded that scoring the first goal of the match can mean victory in over 80% of cases in women's football if the team is well positioned in the ranking.[3]

Another study[15] that has analyzed women's football from a perspective of space management and the duels that occur in it concluded that a quarter of the goals are produced via a center, and with a predominance of 1×1 offensive duels.[9]

Given these data, it is possible to confirm that attacking in women's football is currently being studied, but available research is scarce and quite inconclusive. Furthermore, there is very little work that has covered the analysis of the offensive process from a predictive and multivariate point of view that would allow theoretical performance models to be transferred to the applied field: trainers and technical bodies.

Therefore, the aim of this study is threefold: on a univariate level to discover and describe how attacks are produced in women's football, differentiating between those which end in success and those which do not; on a bivariate level to identify the criteria that have a statistically significant relationship with success (goal, shot on goal or pass into the area); and on a multivariate level, two predictive analyses were carried out (logistical regression and decision tree), that allowed us to propose theoretical models of offensive success. It is hypothesized that the time of possession in the rival midfield (MO) and the tactical intention are criteria associated with success (goal, shot, or sending to the penalty area).

2 | METHOD

2.1 | Design

The observational design used was nomothetic, intersessional, and multidimensional follow-up.[16,17] The systematic observation carried out was non-participant and active, using an observational sampling "all occurrence".

2.2 | Participants

A total of 6063 attacks starting in possession of the ball have been collected and analyzed during 68 matches held during the last two FIFA Women's World Cups, 2015 (Canada) and 2019 (France). This represents an average of 89.2 attacks per game. The inclusion criteria used for the recording of offensive actions were adapted from Garganta.[18] Actions in which the attacking team fulfilled any of the following requirements were coded: (i) three consecutive contacts with the ball or (ii) a finished pass—as long as it lasted more than 3 s, or (iii) a shot taken. The offensive actions lasted from the first contact with the ball up to (i) possession changing to the rival team or (ii) there being a regulatory interruption in the game.

The principles of invasion sports are established as a relationship between the two teams, where each one must coordinate their actions to recover, keep, and send the ball to the end zone.[19] Following the contributions of Wade,[20] when a team has possession of the ball, each player of the team is on attack, with the aim of maintaining possession of the ball and advancing the ball toward the scoring opportunity. Wade's proposal on the attack has been collated and used in subsequent studies.[21–24]

Matches were recorded from public images broadcasted on television, and through a post-event record, thus ensuring respect for behavior spontaneity, as well as the registration in its natural environment. According to the Belmont Report,[25] the use of public images for research purposes does not require consent.

2.3 | Instruments

The observation instrument proposed by Maneiro et al.[10] has been used (Table S1). This instrument is well suited to the object of study and has been used in similar investigations.[10–12] Altmann's[26] advice on molar and molecular fit of behaviors and categories has been followed. To achieve this balance, each behavior unit conforms to the rule of the three Ds: delimitable (identifiable by themselves, presenting their own identity); denominable (the adornment of a name allows its distinction); and definable (it maintains the basis for its operation in the subsequent evaluation) (delimitable, denominable, and definable).[27,28]

2.4 | Procedure

Prior to the coding process, and to reduce interobserver variability, eight training sessions have been carried out, following Anguera et al.[29] and Manolov et al.[30] In the first place, four observers were selected for data collection, three of them are PhDs in Sports Sciences, who are also national soccer coaches, and with experience in studies of this type. To ensure methodological quality, one of the co-authors is an expert in observational methodology, who has ensured that all the methodological steps are correct. Secondly, the training sessions have been 2 h long each. The first three sessions were carried out in groups with the selected observers. The theoretical approach of this study was presented to them, the behaviors to be observed were delimited, the observation instrument was exposed, and the observers were trained in the use of the Lince Plus recording instrument.[31] The fourth session consisted of the observation and recording by the observers of 20 offensive actions previously selected by the principal investigator, ordered from least to greatest complexity. Once the actions were recorded by those observed, the discrepancies found were discussed. The fifth and sixth sessions were carried out individually with each of the observed. The delimitation of the recorded actions was carried out previously by the principal investigator, and those observed were instructed in the recording of the actions. The last two sessions were also carried out individually, and in them, the concordance coefficient of Cohen's Kappa[32] was verified between the principal investigator and each of the observed. Finally, two files were given to each of the observers with the offensive actions under analysis. The actions used to obtain the value of the coefficient of agreement represented 10.8% of the total actions ($n = 654$). The analysis of said actions was carried out individually and sent to the principal investigator of the study.

Data quality control was carried out using the IBM SPSS Statistics 25 program by means of an interobserver concordance analysis by Cohen's Kappa coefficient[32] for each of the criteria, the overall value being very good (0.83) (Table S2) according to the scales of Fleiss, Levin, and Paik.[33]

2.5 | Statistical analysis

R programming software has been used, using the libraries "compareGroups,[34]" "rpart,[35]" and "partykit[36]". A recategorization of the variable "move outcome" was performed, which consisted of transforming its four categories—goal, shot, sending to the area and no success—into two categories: the category NO SUCCESS (NE), which already existed, and the SUCCESS (E) category, which included the other three. A descriptive analysis of all the variables was carried out at the same time as a bivariate study between the explained variable "move outcome" and the rest of the variables that make up the observation instrument. Predictor criterion showed a significant relationship with the explained variable. This information was subsequently used to configure a logistic regression model.

As a final analysis, a multivariable technique based on decision trees was incorporated to know what the probabilities of NO SUCCESS and SUCCESS were in the different combinations (Figure 1). It is a nonparametric approach, that is, without supposed distributions. It has an easy control of lost values and strongly asymmetric data without the need to resort to data transformation. It is a robust analysis of outliers and, in addition, allows the analysis of sequential decisions based on the use of associated probabilities. For this, the chi-square automatic interaction detector (CHAID) was used as a growth method, which consists of a statistical and multidirectional tree algorithm that scans data quickly and efficiently, and creates segments and profiles compared to the desired result. In each step, CHAID chooses the predictor variable that presents the strongest interaction with the explained variable. The categories of each predictor merge if they are not significantly different from the predictive variable.

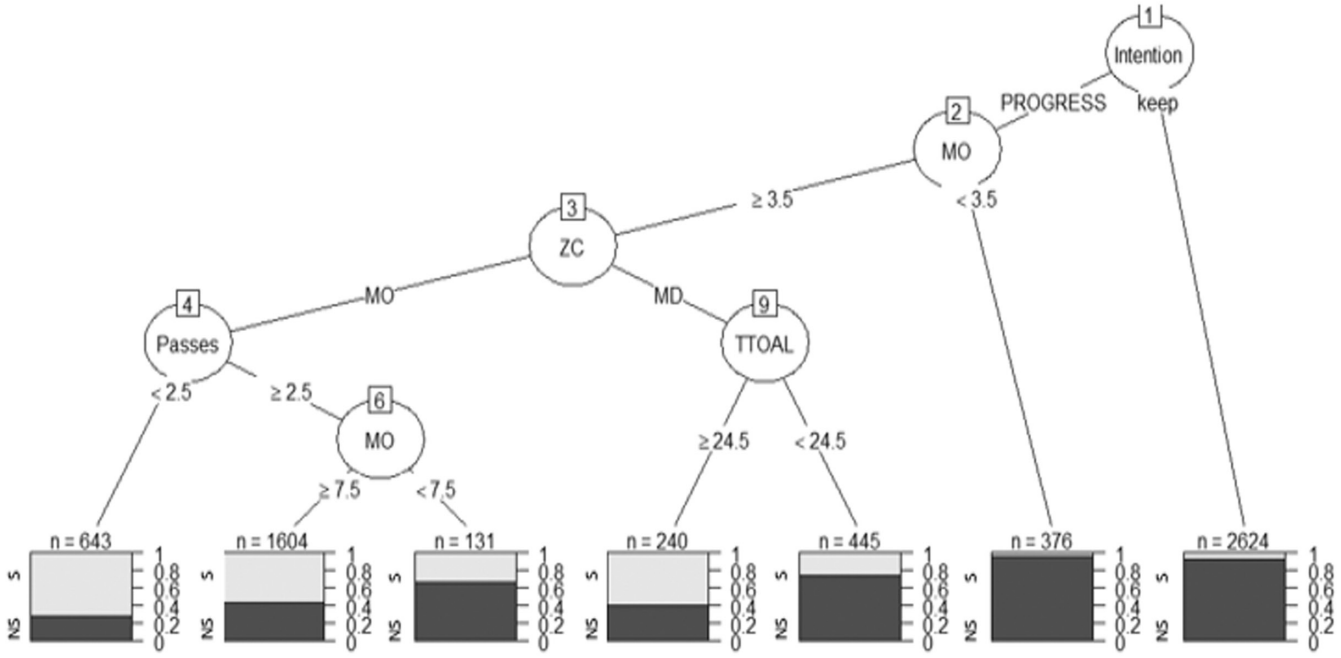

FIGURE 1 Decision tree representation

TABLE 1 Summary descriptives table by groups of "move outcome"

	Success ($n = 1895$)	No success ($n = 4198$)	p Value
Competition			<0.001
FIFA Women's World Cup 2019	579 (30.6%)	1744 (41.8%)	
FIFA Women's World Cup 2015	1316 (69.4%)	2424 (58.2%)	
Half-time:			<0.001
FirstHalf	884 (46.6%)	2288 (54.9%)	
SecondHalf	1011 (53.4%)	1880 (45.1%)	
Start-form			0.081
Set-play (estatic)	525 (27.7%)	1248 (29.9%)	
Transition (dynamic)	1370 (72.3%)	2920 (70.1%)	
Intention			<0.001
Keep	217 (11.5%)	2407 (57.7%)	
Progress	1678 (88.5%)	1761 (42.3%)	
MO	11.0 [8.00; 15.0]	5.00 [1.00; 10.0]	<0.001
ZC			<0.001
1 (MD)	306 (16.1%)	2588 (62.1%)	
2 (MO)	1589 (83.9%)	1580 (37.9%)	
Match-status			0.028
Winning	547 (28.9%)	1068 (25.6%)	
Drawing	836 (44,1%)	1905 (45.7%)	
Losing	512 (27.0%)	1195 (28.7%)	

3 | RESULTS

A total of 6063 ball possessions were collated and analyzed from the FIFA Women's World Cup 2015 and 2019. An initial univariate analysis of the final success of these actions revealed that just 2.1% of possessions ended in a goal, 11.2% ended with a shot on goal, 17.9% ended with a pass into the penalty area, and practically 69% of all possessions in women's elite football ended unsuccessfully.

Table 1 below shows the results for the categorical criteria. The frequencies were calculated in each category and the percentage with respect to the total. The significance level was fixed at $p < 0.05$. A chi-square test was done to find out the explanatory criteria that had a significant relationship with the explained criterion and were "Half time," "Start-form," "Intention," "MO," "ZC," and "Match Status."

The values *p. overall* equate to the values *p. chisq*. All the explanatory criteria maintain a significant relationship with the explained variable *MoveOutcome*.

The ROC test was used to discover the significance of the quantitative criteria, analyzing the area below the ROC curve and taking the values above 0.6 as significant. There were four quantitative criteria: MD, MO, Total time, and Passes. Only the MO criterion has been statistically significant (0.756).

3.1 | Results of the logistic regression analysis

For the qualitative criteria, we studied the significant relationships of the explanatory criteria with the explained criterion "*MoveOutcome*." A logistic regression test was carried out to know which was the most suitable model. The formula for the selected model is as follows:

$$\text{MoveOutcome} - > \text{HalfTime} + \text{Intention} + \text{MD} + \text{MO} + \text{ZC} + \text{Passes} + \text{MatchStatus} + \text{Competition}$$

Regarding the values of the deviance, the median is close to 0 (0.2716) and shows an almost perfect symmetry (Deviance residuals: Min = −2.60; 1Q = −0.95; 3Q = 0.69; Max = 2.21). Table 2 shows the coefficients.

The "*backward*" strategy was used to select the model in such a way that the initial model was subjected to a gradual degradation of explanatory criteria until the differences between the models in the adjustment were not significant.

The resulting model from this process came out as (Table 3):

$$(\text{MoveOutcome}) \sim \text{HalfTime} + \text{StartForm} + \text{Intention} + \text{MO} + \text{ZC}$$

In this model, all the explanatory criteria are significant with the explained variable "*MoveOutcome*" except the variable StartFormDYNAMIC.

Regarding the values of the deviance, the median is close to 0 (0.2739) and shows an almost perfect symmetry (Deviance residuals: Min = −2.59; 1Q = −0.95; 3Q = 0.69; Max = 2.15).

3.1.1 | Coefficients

We are interested in the sign of the estimated values. We observed that all the criteria have negative coefficients except one. Playing in the second half HalfTimeSECOND HALF, when the intention is *progress* (IntentionPROGRESS), MO, and ZCMO decrease the possibility of success. However, the variable *Intercept* that is the large mean factor and StartFormDYNAMIC facilitate success. Observing the value of the effect size (z value) those that have most influence are firstly "IntentionPROGRESS" followed by "ZCMO" (Table 4).

When the ODD ratio values are greater than 1, there is a positive association; lower than 1 means the association is negative. When they are the same as 0, it means they are equal (Table 5).

TABLE 2 Coefficient table

	Estimate Std.	Error	Z value	Pr (>| z |)
(Intercept)	3.354468	0.129614	25.882	<2e-16***
HalfTime Second Half	−0.193557	0.069662	−2.779	0.00546**
StartForm dynamic	0.121563	0.074021	1.642	0.10053
Intention progress	−1.951520	0.087776	−22.233	<2e-16***
MO	−0.048902	0.005568	−8.782	<2e-16***
ZC MO	−1.307935	0.083717	−15.623	<2e-16***
MatchStatus drawing	0.011242	0.082326	0.137	0.89139
MatchStatus losing	0.110908	0.088791	1.249	0.21163
Competition FIFA Women's World Cup15	0.097049	0.074972	1.294	0.19551

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

The variable IntentionPROGRESS decreases by 0.1467525 the possibility of achieving success against IntentionKEEP. With respect to the variable ZCMO, it has a lower probability of 0.2659574 of achieving success with respect to ZCMD. With respect to the variable HalfTimeSECOND HALF, it has a probability of 0.8374415 times of achieving success than in HalfTimeFIRSTHALF. With respect to the variable MO, it has 0.9536981 times the probability of not achieving success against MD. The variable StartFormDYNAMIC increases by 1.1226997 times the probability of achieving success in relation to StartFormSTATIC. This variable turned out not to be significant in the model.

The goodness-of-fit tests prove whether the proposed model fits the data. We have various options: McFadden's pseudo-R squared 0.2563223 and Nagelkerke 0.3834014 (coefficient of determination is 38.34%); Cox and Snell's with a value of 0.2727009 to check whether the proposed model can explain what is observed. These values indicate the variance part of the dependent variable explained by the model. The R^2 of the model tends to be between the Snell value and Nagelkerke. Finally, the Hosmer and Lemenshow test with a value of 0.3438896 is not significant, so the proposed model fits the data.

The distribution function is calculated on the basis of the model's prediction:

$$\pi_{(\text{MoveOutcome})} = \frac{\exp^{(3.42\ -1.92\ -1.32\ -0.17\ -0.05)}}{1+\exp^{(3.42\ -1.92\ -1.32\ -0.17\ -0.05)}} = \frac{0.05}{0.95} = 5.2631\ (52.6\%)$$

TABLE 3 Results of the "backward" strategy

	Df	Deviance	ATC
<None>		5604.5	5616.5
StartForm	1	5606.9	5616.9
HalfTime	1	5611.7	5621.7
MO	1	5685.5	5695.7
ZC	1	5876.2	5886.2
Intention	1	6246.3	6256.3

3.2 | Results of the decision tree analysis

A total of 6063 observations were carried out and 1895 were lost. The majority were "No success" (68.74%) while 31.25% were "Success" (Table 6).

Due to the large size of the tree, of the 7 terminal nodes only those with notable differences will be mentioned. In the case of Intention = keep with 2624 observations and lost values, the most frequent was "No Success" with 91.73% and 8.26% "Success," formulating the first terminal node. For the variable N≥4, the format of "no Success" with 376 observed and 15 lost has a 3% probability of "Success" and 96.01% "No success" consolidating a second terminal node. Another terminal node comes from the configuration TToal≥25 which in the case of "No Success" has a probability of 74.15%. A following differential node is the variable MO≥8 with a probability of 67.17% of "No Success" in 131 observations and 43 losses. Finally, in the variable "Passes < 3" the "Success" has a probability of 71.22% node, split, n, loss, yval, (yprob). The representation of the decision tree can be seen in Figure 1.

4 | DISCUSSION

This work arose with a triple aim: on the one hand to analyze how ball possessions are produced in women's elite football in terms of incidence, efficiency, and habitual practices; on the contrary, to identify the criteria that have a statistically significant relationship with success; and also to propose predictive models of success that can provide solutions to the applied field and improve player performance.

The available results allow us to confirm the hypotheses proposed in this study. The results of the bivariate and multivariate analyzes confirm that the time of possession in the rival field (MO) and the tactical intention are criteria associated with success.

TABLE 4 Coefficient table

	Estimate Std.	Error	Z value	Pr (>| z |)
(Intercept)	3.420743	0.109248	31.312	<2e-16***
Intention progress	−1.919008	0.083779	−22.906	<2e-16***
ZC MO	−1.324419	0.082738	−16.007	<2e-16***
HalfTime SecondHalf	0.177404	0.065752	−2.698	0.00697
MO	−0.047408	0.005351	−8.859	<2e-16***
StartForm dynamic	0.115736	0.073867	1.567	0.11716

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

TABLE 5 Odd ratio values

	OR	2.5%	97.5%
(Intercept)	30.5921425	24.7720042	38.0177842
Intention progress	0.1467525	0.1242810	0.1726196
ZC MO	0.2659574	0.2259418	0.3125252
HalfTime SecondHalf	0.8374415	0.7361552	0.9526210
MO	0.9536981	0.9436925	0.9637032
StartForm dynamic	1.1226997	0.9712882	1.2975490

TABLE 6 Theoretical representation of the model

1) root 6063 1895 No success (0.31255154 0.68744846)
2) Intention = PROGRESS 3439 1678 No success (0.48793254 0.51206746)
4) MO > =3.53063 1400 Success (0.54293177 0.45706823)
8) ZC = MO 2378 973 Success (0.59083263 0.40916737)
16) Passes<2. 643 185 Success (0.71228616 0.28771384) *
17) Passes> = 2.51735 788 Success (0.54582133 0.45417867)
34) MO > =7.51604 700 Success (0.56359102 0.43640898) *
35) MO <7.5131 43 No success (0.32824427 0.67175573) *
9) ZC = MD 685258 No success (0.37664234 0.62335766)
18) TTOAL> = 24.5240 97 Success (0.59583333 0.40416667) *
19) TTOAL < 24. 445 115 No success (0.25842697 0.74157303) *
5) MO < 3.5376 15 No success (0.03989362 0.96010638) *
3) Intention = keep 2624 217 No success (0.08269817 0.91730183) *
17) Passes> = 2.51735 788 Success (0.54582133 0.45417867)
34) MO > = 7.51604 700 Success (0.56359102 0.43640898) *
35) MO < 7.5131 43 No success (0.32824427 0.67175573) *
9) ZC = MD 685258 No success (0.37664234 0.62335766)
18) TTOAL > = 24.5240 97 Success (0.59583333 0.40416667) *
19) TTOAL < 24. 445 115 No success (0.25842697 0.74157303) *
5) MO < 3.5376 15 No success (0.03989362 0.96010638) *
3) Intention = keep 2624 217 No success (0.08269817 0.91730183)*

Note: * The terminal nodes of the decision tree, and the R program exports that way.

Regarding the effectiveness, it must be highlighted that success (goal, shot on goal or pass into the area) is a very difficult thing to achieve in female football. To be precise, 7 out of every 10 attacks which start with ball possession end unsuccessfully. What's more, almost 6 possessions are required to achieve a pass into the penalty area, nearly 9 to achieve a shot on goal, and almost 48 possessions to achieve a goal. These data corroborate a previous work about women's football by Iván-Baragaño et al.,[11] and we also find values similar to those of men's football.[37–39] Studies such as the one by Shafizadeh et al.[40] state that this inefficiency may be due to technical aspects, poor decision-making, low level of concentration, or the high speed of these actions, although it should also be considered the defensive success of the rival.[41]

At the bivariate level, there are significant differences between the successful attacks that are made between the first and second part of the match. Specifically, more successful attacks (goal, shot on goal, or delivery into the area) are made in the second half. A possible explanation could be the tiredness of the rival team, which reduces their tactical performance, or the different styles of play.[23,42] On the contrary, the teams that initiate the attack through a transition (recovery of the ball during the game) are more successful than those that initiate through regulatory incidence, behavior that corroborates the previous work of Kirkendall et al.,[43] who observe that 62% of attacks begin after possession of the ball in women's football. The data in men's football still require more scientific literature on this aspect.[44] Another of the criteria that has remained significant has been the tactical intention to progress once the ball is recovered. This result confirms previous findings in men's soccer[42] and also in women's football,[12] on the advantage of taking advantage of moments of change of ball possession to attack the rival defense.

On the contrary, a statistically significant relationship has been found between the success of the attack and the time of possession, with possession in the opposite field (offensive zone of the observed team) being one of the variables that explain success. This result has been previously compared both in men's football in general,[45] and also in relation to the partial result in particular,[23] as well as in women's football.[12] In general, the best teams have the ball in the offensive zone, close to the rival goal, with the aim of seeking a finish.

Lastly, match status is also a criterion that significantly modulates team attacks. This criterion has been extensively studied in male soccer,[24,46–48] although in women's football has not yet been addressed in depth.[10] Specifically, although almost half of the attacks that end in success (goal, shot, or shot) occur with the match status of a draw, a higher percentage is observed when they are winning on the scoreboard than when they are losing. One of the possible explanations may be the difference in quality between the teams,[24,49] where there is a notable difference between the best and the worst teams, which means that the teams that are winning attack more successfully (finalization in a goal, in a shot or in a shipment to the area), than the teams that are losing (they attack more, but with less offensive success).

The predictive multivariate results are conclusive: on a tactical-strategic level, women's football teams obtain better results when they make rapid attacks, with short ball possessions, preferably in the opponent's half and

when looking for a quick finish. However, other criteria such as match status (winning, drawing, losing) or the interaction context (which line recuperates the ball against which rival line) are not criteria that have shown information gain and were not significant. With the proposed predictive model of logistic regression, the teams would go from a 31% success rate (in the absence of a model), to 52.6%, thus increasing by 21.6% their options of achieving a goal, a shot on goal or a pass into the area. On the contrary, and with the aim of complementing and confirming these results, another multivariate technique was implemented—the decision tree (Figure 1) and again shows the alternative of the teams that carry out rapid attacks, keeping possession in the rival midfield and with a reduced number of passes. Specifically, the nodes "intention," "MO," and "passes" are those which show a higher information gain with respect to success, with practically 70% of success with $n = 643$. However, the teams that carry out slow attacks, with long possession times, only achieve success in 8.2% of their attacks ($n = 2624$).

4.1 | Perspective

In the present investigation, a general description of how the attack occurs in high-level women's football is offered. In general terms, the offensive process in women's football is clearly ineffective, to the extent that only 2.1% of attacks end in a goal and practically 70% end without a shot on goal or without the ball entering the area rival. On the contrary, it is possible to affirm that there are a series of criteria that are related in a statistically significant way to success in the attack, such as carrying out the attack in the second half, recovering the ball dynamically and not through regulation, with an intention to progress toward the rival goal and possess the ball in the offensive sector of the field. The theoretical model presented makes it possible to explain 52.6% of offensive success, information that teams could use to optimize and do a more effective job in attacking training, and thus improve their performance during the competition. These findings may have a direct transfer to the applied field, to the extent that empirically supported recommendations are offered to improve sports performance. These studies allow complementing other studies of a physical[50] or physiological[51] nature.

5 | CONCLUSION

There are three main conclusions that can be extrapolated from this work: Firstly, the offensive phase or attack in women's football is ineffective, to the extent that 7 out of every 10 possessions do not even reach the opponent's area; secondly, the criteria that show a statistically significant relationship with success are as follows: "half-time," "start-form," "intention," "MO," and ZC". Lastly, the proposed multivariate models indicate that the most effective way of achieving offensive success is via rapid attacks based on short ball possessions in the opponent's half with the intervention of few players.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of a Spanish government subprojects *Mixed method approach on performance analysis (in training and competition) in elite and academy sport* [PGC2018-098742-B-C33] and *Integration ways between qualitative and quantitative data, multiple case development, and synthesis review as main axis for an innovative future in physical activity and sports research* [PGC2018-098742-B-C31] (Ministerio de Ciencia, Innovación y Universidades, Programa Estatal de Generación de Conocimiento y Fortalecimiento Científico y Tecnológico del Sistema I+D+i), that is part of the coordinated project *New approach of research in physical activity and sport from mixed methods perspective* (NARPAS_MM) [SPGC201800X098742CV0].

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Rubén Maneiro https://orcid.org/0000-0002-2215-9720
Iyán Iván-Baragaño https://orcid.org/0000-0002-7136-4819
José L. Losada https://orcid.org/0000-0001-6134-1819
Antonio Ardá https://orcid.org/0000-0002-0285-2149

REFERENCES

1. Lago I, Lago-Peñas S, Lago-Peñas C. Waiting or acting? The gender gap in international football success. *Int Rev Sport Sociol*. 2021;101269022110607. doi:10.1177/10126902211060727
2. Okholm K, Wang A, Mehta R, et al. Research on women's football: a scoping review. *Sci Med Footb*. 2021;1-10. doi:10.1080/24733938.2020.1868560
3. Kirkendall DT. Issues in training the female player. *Br J Sports Med*. 2007;41:64-67. doi:10.1136/bjsm.2007.036970
4. de Jong LM, Gastin PB, Angelova M, et al. Technical determinants of success in professional women's soccer: a wider range of variables reveals new insights. *PLoS One*. 2020;15(10):e0240992.
5. Kirkendall DT, Krustrup P. Studying professional and recreational female footballers: a bibliometric exercise. *Scand J Med Sci Sports*. 2022;32:12-26.
6. Mombaerts E. *Fútbol. Del análisis del juego a la formación del jugador [Soccer. From game analysis to player training]*. INDE; 2000.

7. Scanlan M, Harms C, Wilkie JC, et al. The creation of goal scoring opportunities at the 2015 women's world cup. *Int J Sports Sci Coach*. 2020;15(5–6):803-808. doi:10.1177/1747954120942051
8. Kubayi A, Larkin P. Technical performance of soccer teams according to match outcome at the 2019 FIFA Women's world cup. *Int J Perform Anal Sport*. 2020;20:908-916. doi:10.1080/24748668.2020.1809320
9. Soroka A, Bergier J. Actions with the ball that determine the effectiveness of play in women's football. *J Hum Kinet*. 2010;26:97-104. doi:10.2478/v10078-010-0053-y
10. Maneiro R, Losada JL, Casal CA, Ardá A. The influence of match status on ball possession in high performance Women's football. *Front Psychol*. 2020;11:487. doi:10.3389/fpsyg.2020.00487
11. Iván-Baragaño I, Maneiro R, Losada JL, Ardá A. Multivariate analysis of the offensive phase in high-performanceWomen's soccer: a mixed methods study. *Sustain For*. 2021;13(11):6379. doi:10.3390/su13116379
12. Maneiro R, Losada JL, Casal C, et al. Identification of explanatory variables in possession of the ball in high-performance Women's football. *Int J Environ Res Public Health*. 2021;18(11):5922. doi:10.3390/ijerph18115922
13. Casal C, Losada JL, Maneiro R, et al. Gender differences in technical-tactical behaviour of La Liga Spanish football teams. *J Hum Sport Exerc*. 2020;16:37-52. doi:10.14198/jhse.2021.161.04
14. Ibáñez S, Pérez-Goye JA, Courel-Ibáñez J, et al. The impact of scoring first on match outcome in women's professional football. *Int J Perform Anal*. 2018;18(2):318-326. doi:10.1080/24748668.2018.1475197
15. Mara J, Wheeler KW, Lyons K. Attacking strategies that lead to goal scoring opportunities in high level women's football. *Int J Sports Sci Coach*. 2012;7(3):565-577. doi:10.1260/1747-9541.7.3.565
16. Anguera MT. Observational typology. *Qual Quant*. 1979;13(6):449-484.
17. Anguera MT, Hernández-Mendo A. Avances en estudios observacionales en Ciencias del Deporte desde los mixed methods [Advances in observational studies in Sports Sciences from mixed methods]. *Cuad Psico Dep*. 2016;16:17-30.
18. Garganta J. Modelaçao táctica do jogo de Futebol. Estudo da organizaçao da fase ofensiva em equipas de alto rendimiento [tactical modeling of the football game. Study of the organization of the offensive phase in high-performance teams] (doctoral thesis). Universidad de Oporto; 1997.
19. Gréhaigne JF, Godbout P. Tactical knowledge in team sports from a constructivist and cognitivist perspective. *Quest*. 1995;47(4):490-505.
20. Wade A. *Principles of Team Play*. Reedswain Inc.; 1996.
21. Castellano J, Casamichana D, Lago C. The use of match statistics that discriminate between successful and unsuccessful soccer teams. *J Hum Kinet*. 2012;31:139-147.
22. Hewitt A, Greenham G, Norton K. Game style in soccer: what is it and can we quantify it? *Int J Perform Anal Sport*. 2016;16(1):355-372.
23. Lago-Peñas C. The influence of match location, quality of opposition, y match status on possession strategies in professional association football. *J Sports Sci*. 2009;27(13):1463-1469.
24. Lago-Peñas C, Martín-Acero R. Determinants of possession of the ball in soccer. *J Sports Sci*. 2007;25(9):969-974.
25. National Commission for the Protection of Human Subjects of Biomedicaland Behavioral Research. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. Superintendent of Documents; National Commission for the Protection of Human Subjects of Biomedicaland Behavioral Research; 1979.
26. Altmann SA. Sociobiology of rhesus monkeys. II: stochastics of social communication. *J Theor Biol*. 1965;8(3):490-522.
27. Anguera MT, Blanco-Villaseñor A, Losada JL, Sánchez-Algarra P, Onwuegbuzie AJ. Revisiting the difference between mixed methods y multimethods: ¿is it all in the name? *Qual Quant*. 2018;52:2757-2770. doi:10.1007/s11135-018-0700-2
28. Anguera MT, Camerino O, Castañer M, Sánchez-Algarra P, Onwuegbuzie AJ. The specificity of observational studies in physical activity y sports sciences: moving forward in mixed methods research y proposals for achieving quantitative y qualitative symmetry. *Front Psychol*. 2017;8:2196. doi:10.3389/fpsyg.2017.02196
29. Anguera MT, Blanco-Villaseñor A, Losada JL, et al. Análisis de la competencia en la selección de observadores [analysis of competition in the selection of observers]. *Met Cienc Comport [Behavl Sci Method]*. 1999;1(1):95-114.
30. Manolov R, Losada JL. Simulation theory applied to direct systematic observation. *Front Psychol*. 2017;8:905. doi:10.3389/fpsyg.2017.00905
31. Soto A, Camerino O, Iglesias X, et al. LINCE PLUS: research software for behaviour video analysis. *Apunts Educación Física y Deportes*. 2019;137:149-153. doi:10.5672/apunts.20140983.es(2019/3).137.11
32. Cohen JA. Coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37-46. doi:10.1177/001316446002000104
33. Fleiss JL, Levin B, Paik MC. *Statistical Methods for Rates and Proportions*. 3rd ed. John Wiley & Sons; 2003.
34. Subirana I, Sanz H, Vila J. Building bivariate tables: the compareGroups package for R. *J Stat Softw*. 2014;57:1-16.
35. Therneau T, Atkinson B. rpart: Recursive partitioning and regression trees. R Package Versión. 201;4:1–15.
36. Zeileis A, Hothorn T, Hornik K. Model-based recursive partitioning. *J Comput Graph Stat*. 2008;17:492-514. doi:10.1198/106186008x319331
37. Wallace JL, Norton KI. Evolution of world cup soccer final games 1966–2010: game structure, speed and play patterns. *J Sci Med Sport*. 2014;17(2):223-228.
38. Brito-Souza D, López-Del Campo R, Blanco-Pita H, et al. A new paradigm to understand success in professional football: analysis of match statistics in LaLiga for 8 complete seasons. *Int J Perform Anal Sport*. 2019;19(4):543-555.
39. Lago-Ballesteros J, Lago-Peñas C, Rey E. The effect of playing tactics and situational variables on achieving score-box possessions in a professional soccer team. *J Sports Sci*. 2012;30:1455-1461. doi:10.1080/02640414.2012.712715
40. Shafizadeh M, Gray S, Sproule J, McMorris T. An exploratory analysis of losing possession in professional soccer. *Int J Perform Anal Sport*. 2012;12(1):14-23.
41. Vogelbein M, Nopp S, Hökelmann A. Defensive transition in soccer–are prompt possession regains a measure of success? A quantitative analysis of German Fußball-Bundesliga 2010/2011. *J Sports Sci*. 2014;32(11):1076-1083.

42. Tenga A, Holme I, Ronglan LT, Bahr R. Effect of playing tactics on goal scoring in Norwegian professional soccer. *J Sports Sci*. 2010;28(3):237-244.
43. Kirkendall DT, Dowd WW, DiCicco AD. Patterns of successful attacks: a comparison of the men's and women's world cup. *R Fútbol Ciencia*. 2002;1:29-36.
44. Sarmento H, Marcelino R, Anguera MT, CampaniÇo J, Matos N, LeitÃo J. Match analysis in football: a systematic review. *J Sports Sci*. 2014;32(20):1831-1843.
45. Casal C, Maneiro R, Ardá T, et al. Possession zone as a performance indicator in football. The game of the best teams. *Front Psychol*. 2017;8:1176.
46. Jones PD, James N, Mellalieu SD. Possession as a performance indicator in soccer. *Int J Perform Anal Sport*. 2004;4:98-102. doi :10.1080/24748668.2004.11868295
47. Lago-Peñas C, Dellal A. Ball possession strategies in elite football according to the evolution of the match-score: the influence of situational variables. *J Hum Kinet*. 2010;25:93-100.
48. Bloomfield JR, Polman RC, O'Donoghue PG. Effects of scoreline on team strategies in FA premier league soccer. *J Sports Sci*. 2005;23:192-193.
49. Castellano J, Blanco-Villaseñor A, Álvarez D. Contextual variables and time-motion analysis in soccer. *Int J Sports Med*. 2011;32(06):415-421.
50. Vescovi JD, Rupf R, Brown TD, Marques MC. Physical performance characteristics of high-level female soccer players 12–21 years of age. *Scand J Med Sci Sports*. 2011;21(5):670-678.
51. Ali A, Gardiner R, Foskett A, Gant N. Fluid balance, thermoregulation and sprint and passing skill performance in female soccer players. *Scand J Med Sci Sports*. 2011;21(3):437-445.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Maneiro R, Iván-Baragaño I, Losada JL, Ardá A. Deciphering the offensive process in women's elite football: A multivariate study. *Scand J Med Sci Sports*. 2022;00:1-10. doi: 10.1111/sms.14206