IS IT POSSIBLE TO ESTIMATE MATCH RESULT IN VOLLEYBALL: A NEW PREDICTION MODEL

Cengiz Akarçeşme

Gazi University School of Physical Education and Sports, Ankara, Turkey

Address for correspondence:
Cengiz Akarçeşme
Gazi University School of Physical Education and Sports
Abant Sokak No 12
Postal Code: 06560 Ankara, Turkey
E-mail: cengizakarcesme@gazi.edu.tr

Abstract This study investigates the power of variables in a logistic regression model (the efficacy model or (EM)) to explain the match results in the Turkish Men's Volleyball League (TMVL) and the Turkish Women's Volleyball League (TWVL) in terms of the players' positions. The dependent variable was the match result, and the power of the variables libero player efficiency (LPE), setter efficiency (SE), middle blocker efficiency (MBE), outside hitter efficiency (OHE) and universal player efficiency (UPE) were separately investigated for both genders. The EM accurately classified 83.45% of the games won and lost in the TWVL. The sensitivity (proportion of won games classified as won) and specificity (proportion of lost games classified as lost) was 85.03 and 81.88%, respectively. In the TMVL analysis, the classification accuracy, sensitivity and specificity were 78.23, 78.77 and 77.70%, respectively. Moreover, for both genders, the match results were chiefly explained by the SE, MBE, OHE and UPE. The LPE variable could not predict the results in the TWVL.

Key words volleyball, gender, logistic regression, result, efficacy model

Introduction

Almost all Olympic sports hold separate tournaments for men and women. The main reason for not competing men against women is the gender factor, which is considered to have an important influence on sport performance (Beim, Winter, 2003, p. 35). Men and women largely differ in their physical body composition, muscle mass, hormonal secretions and oxygen consumption (Rickenlund et al., 2003; Korhonen, Mero, Suominen, 2003). Kinanthropometric properties significantly differ by gender and somatotype, affecting the selected sports, positions and performances of men and women (Gualdi-Russo, Graziani, 1993). However, the only gender-based rule of volleyball is the height of the net. The similarities and differences between male and female volleyball players have not been investigated from a modelling perspective. In volleyball, as in other branches of sports, numerous statistical data have been collected by match analysis programs. To interpret this large body of data and make inferences, we need statistical tools. In particular, statistical models can reveal the relationships among various variables and extract information

on team performance. Such information can assist team coaches in developing their training strategies and game tactics. Rendering statistical information interpretable will also increase the accuracy and effectiveness of decision-making processes. Among the acquired data, a coach can identify elements that encourage winning approaches by the team or set numerical targets that are likely to be reached. At this stage, the biggest problem is the dozens of volleyball game variations that may affect basic techniques such as pass, attack, block, serve and serve reception and combinations of these variations. In addition, considering the team elements and competition against the other team, the effects of individual elements on the game cannot be identified.

The variables affecting the results of volleyball matches have been analysed in the literature, but the effects of movements used during the game, referred to as technical elements, have been rarely investigated. Asterios, Kostantinos, Athanasios, Dimitrios (2009) investigated the relationship between some of the technical elements and the match results of the Men's Volleyball World Championship, held in Japan in 2006. Using discriminant analysis, they found that attack error, jump serve point, quick ball error and jump serves effectively explained the match results. Monteiro, Mesquita, Marcelino (2009) studied the relationship between set outcome and the dig and attack efficacy in female volleyball. They reported a significant correlation between attack efficacy and set outcome, but no correlation between dig efficacy and set outcome. However, attack and dig efficacy were interdependent. Araujo, Mesquita, Marcelino (2009) observed a high correlation between the set blocking system and set outcome. Drikos, Kountouris, Laios, Laios (2009) calculated the efficacy variables of volleyball techniques employed by Greek premier league male volleyball teams during 2005–2006 and determined that the team performance was most affected by the serve and attack efficacy ratios. Marelic, Resetar, Jankovic (2004) studied the effects of technical elements on the match result and proposed discriminant functions for predicting the match outcome from the serve, serve reception, spike in the phase of attack, block and spike in the phase of counter attack. They concluded that all these variables, especially the spike in the attack phase, significantly discriminate the win and loss outcomes.

In almost all previous studies on variables affecting match performance, the efficacy variables were either defined by their direct values or expressed as percentages. In the present study, the variable definitions are based on the players' positions. To facilitate their efficacy evaluation, all variables except the *libero player efficiency* (LPE) are calculated as the differences between the scores achieved and lost. In addition, most of the published studies have correlated the variables and identified the effective ones. The present study adopts a predictive logistic regression model called the efficacy model (EM) to extract further information and perform tactical simulations. We also investigate the effects of gender differences on the volleyball variables, which have not been previously reported.

To this end, we evaluate the power of the EM variables to explain the match results in the Turkish Women's Volleyball League (TWVL) and the Turkish Men's Volleyball League (TMVL).

Methods

The efficacy model (EM) employed in this study was developed in Akarçeşme's (2010) PhD thesis. The EM has successfully predicted the match results in women's volleyball. From the models formed in this study Efficiency Model estimates the match result right with 87.65%; Technical Elements Model estimates the match result right with 84.66%.

The EM model outputs a binary variable *result*. Therefore, we employed a logistic link function and analysed the data by logistic regression. Before obtaining the EM, we trialled many variables that would likely affect the match

result. The match outcome did not singly depend on excellent serve reception by the libero, the ratio of excellent serve reception of the team (excepting the libero), points achieved from block, serve and other motions by the setter (for example setter tip, setter kill block), serve reception fault, points from attack and other motions by the outside hitter, points from block, block error and other motions by the middle blocker or the total number of points from attack, attack error and other motions by the universal player. Therefore, we calculated the net differences in player positions between the teams' successes and failures, defined new variables and examined whether these variables significantly explained the match results. Because these derived variables define the efficacies in terms of player positions, they are referred to as *efficacy variables*.

Using the models yielded by the analyses, we predicted the match results of TWVL and TMVL by estimating the model coefficients and conducting a diagnostic study. The initial model prediction was based on data from the women's league. The variance inflation factor (VIF), conditional eigenvalues and index values were derived from the multicollinearity between independent variables. Moreover, the goodness of fit of the model was determined from the Bayesian and Akaike information criteria (BIC and AIC respectively), which quantify the model fitness. The predictive power of the model was assessed by constructing a classification table. Finally, the effect of potentially anomalous data on the goodness of fit was analysed using standardised Pearson residuals and deviance residual statistics. The influence of likely anomalous observations on the fitted function was checked by calculating leverage values. The model based on these new efficacy variables obtained very satisfactory goodness of fit values, confirming the statistical significance of the model. The power of the individual efficacy variables in explaining the match result was significant at the 0.05 level. Accordingly, the variable *libero efficacy* (LE) made the largest contribution to a match win bet (the probability ratio of a match win to a match loss). The other variables also contributed, but to a lesser extent. However, the sensitivity of the coefficient prediction (proportion of won games classified as won) was lower for LE than for the other parameters, because the standard error of parameter prediction was highest for LE.

The study population was the match data of 12 teams competing in Men's and Women's Premier Volleyball Leagues in Turkey during the 2010–2011 season. The sample sizes of the men's and women's leagues were 296 observations (games) in 148 matches and 300 observations in 150 matches, respectively. As these samples covered the whole population, the modelling created from the obtained data was assumed to be valid. The statistical data were used under the consent of the Turkish Volleyball Federation. Moreover, in the leagues of both genders, the teams were entitled to align three foreign players in each game.

In this study, the results of the volleyball matches were correlated with the variables establishing the results for each gender.

Model variables

Result: This variable defines the result of a match. The winning team is represented by 1 and the losing team by 0.

Libero player efficiency (LPE): This variable evaluates the error made by a libero player during serve reception throughout the match. The libero is specialised in defensive skills and cannot attack the ball, so this value starts from 0 and becomes more negative with increasing error. The LPE is generally negative because libero players lack technical flexibility, which would reverse their likelihood of making an error during serve reception.

Setter efficiency (SE): This variable represents the difference between the points scored and points conceded by a setter during all games attended by that setter throughout a match.

Middle blocker efficiency (MBE): The MBE represents the difference between the points scored and points conceded by middle players during all game events throughout a match. The simultaneous evaluation of the MBEs of two middle players is called the MBE.

Outside hitter efficiency (OHE): The OHE represents the difference between the points scored and points conceded by both number 4 hitters during all games throughout a match.

Universal player efficiency (UPE): This variable represents the difference between the points scored and points conceded by a universal player, also referred to as an opposite or power spiker, during all games throughout a match.

In the first part of this study, we constructed the EM, a logistic regression function. For this purpose, we examined the multicollinearity, conducted goodness of fit tests and analysed the influential observations and regression diagnostics (Weisberg, 2005, pp. 211–216). Potential problems with the model were also examined. The efficient observation analysis, which identifies potential anomalous observations that significantly affect the regression model, was conducted first. Certain observations in the dataset may deviate from the general trend and show characteristics that are very distinct from those of other observations. Such anomalous observations can potentially, but do not necessarily, worsen the goodness of fit of the predicted regression model. Efficient observations are anomalous observations that largely affect the predicted parameters (Hosmer, Lemeshow, 2000, pp. 48–56).

One way to handle efficient observations is to exclude them from the analysis. As the predictive model is an average model, such observations attract the regression line, distorting the overall trend. Therefore, we identified and excluded the efficient observations from the model prior to the prediction.

Efficient observations may be analysed by a variety of methods. Any single one of these analyses would give an inaccurate assessment, so several analyses were performed.

The dependent variable *Result* in the EM, which takes values of 1 or 0, is a categorical variable with a binomial distribution. Therefore, the model was based on a logistic function. To explain the dependent variable *Result* through the independent variables LPE, SE, MBE, OHE and UPE, we formulated the EM as follows.

Formula 1

$$\Pi(X) = \frac{e^{\beta 1 + \beta 2 LPE + \beta 3 SE + \beta 4 MBE + \beta 5 UPE + \beta 6 OHE}}{1 + e^{\beta 1 + \beta 2 LPE + \beta 3 SE + \beta 4 MBE + \beta 5 UPE + \beta 6 OHE}}.$$

The logistic regression model is a probability model and $\Pi(X)$ refers to the probability of winning a match. Such a nonlinear function cannot be predicted by the least squares (LS) method, as in linear regression models. To allow an LS interpretation, we perform a logit transformation of $\Pi(X)$. The transformed function is referred to as a logit function.

Formula 2

$$Y = \ln\left(\frac{\Pi(X)}{1 - \Pi(X)}\right) = \beta 1 + \beta 2 LPE + \beta 3 SE + \beta 4 MBE + \beta 5 UPE + \beta 6 OHE.$$

Here, $\left(\dfrac{\Pi(X)}{1 - \Pi(X)}\right)$ defines the bet ratio of winning the match. The βparameters define the variation in the log (bet ratio) for each unit change in the respective efficacy variable.

Results

Table 1 lists the outputs of the predictive logit function.

Table 1. Results of the logistic regression prediction model for Turkish Women's Volleyball League

			Number of obs	=	298	
			LR chi^2	=	175.66	
			Prob > chi^2	=	0.0000	
Log likelihood = −118.72549			Pseudo R^2	=	0.4252	
Result	Odds ratio	Std. err.	Z	P > |z|	[95% Conf. interval]	
LPE	1.0959980	0.1088412	0.92	0.356	0.9021496	1.3314980
SE	1.1471020	0.0576021	2.73	0.006	1.0395820	1.2657430
MBE	1.1712270	0.0322284	5.74	0.000	1.1097340	1.2361280
OHE	1.1648600	0.0266240	6.49	0.000	1.1094590	1.2138590
UPE	1.1504290	0.0326458	4.94	0.000	1.0881920	1.2162270
_cons	0.1435445	0.0531523	−5.24	0.000	0.0694704	0.2966016

As the results of the prediction model, without the setter efficiency (SE) variable, all other variables explains the result variable at the significance level of α = 0.05. The values of the variables effects the result of the match at significance level. The SE variable is significant at the level of α = 0.06.

Table 2 summarises the anomalous value analysis of the prediction model. For this purpose, we calculated the Pregibon residuals, the standardised Pearson residuals and the deviance statistics (Pregibon, 1981). Observations 145 and 286 satisfied the efficient observation criterion for all three statistics, so they were identified to avoid overestimation.

Table 2. Analysis of anomalous observations for the Turkish Women's Volleyball League

Obs	Stdres	p	result	LPE	SE	MBE	OHE	UPE
196	−13.44465	0.9944860	0	−1	1	14	13	21
173	−8.770069	0.9871173	0	−1	1	20	15	6
53	−6.335862	0.9755138	0	−4	4	14	17	5
272	−4.700630	0.9563127	0	−1	−6	12	14	14
157	−4.518397	0.9526947	0	0	3	10	1	20
125	2.226154	0.1717495	1	−1	1	−3	11	−6
293	2.444222	0.1495814	1	−6	−7	−5	14	3
286	2.989338	0.1027186	1	−3	−6	12	−5	−2
58	3.710107	0.0686178	1	−1	1	−8	−1	5
145	6.597115	0.0225887	1	−3	0	−4	−9	3

As shown in Table 2, the teams in observations 196, 173, 53, 272 and 157 lost the matches they were expected to win; conversely, the teams in observations 125, 293, 286, 58 and 145 won the matches they were expected to lose.

Table 3. Analysis of efficient observations 145 and 286 in the Turkish Women's Volleyball League

obs	Pgbon	stdres	dv
145	0.2535106	6.597115	2.753291
286	0.2054256	2.989338	2.133430

Table 3 analyses the anomalous observations 145 and 286 in the Turkish Women's Volleyball League. The Pgbon, stress and dv columns list the Pregibon residuals, Pearson residuals and deviance residuals, respectively. The Pregibon residuals of both observations exceed 0.2 and the Pearson and deviance residuals are higher than 2, confirming that observations 145 and 286 are efficient observations.

After removing these observations and re-performing the logit regression, we obtained the results of Table 4.

Table 4. Results of the logit regression model in the Turkish Women's Volleyball League after removing the efficient observations (observation 145 and 286)

				Number of obs	=	296
				LR chi^2	=	185.68
				Prob > chi^2	=	0.0000
Log likelihood = −112.32597				Pseudo R^2	=	0.4525
Result	Odds ratio	Std. err.	Z	P > |z|	[95% Conf. interval]	
LPE	1.122077	0.1146316	1.13	0.260	0.9184677	1.3708240
SE	1.159955	0.0600320	2.87	0.004	1.0480650	1.2837900
MBE	1.178630	0.0336398	5.76	0.000	1.1145070	1.2464410
OHE	1.178919	0.0289830	6.70	0.000	1.1234610	1.2371160
UPE	1.164912	0.3474270	5.12	0.000	1.0987690	1.2350360
_cons	0.119661	0.0469490	−5.41	0.000	0.0554604	0.2581795

As a final diagnostic analysis of the accuracy of the model, we determined whether any of the independent variables (LE, PE, OOE, PCE and number 4 efficiency (N4E)) were multicollinear. Any linear correlations between these variables would expand the variance of the coefficient predictions, causing errors in the model parameter predictions. Multiple correlations can be identified by several approaches. The correlation matrix provides an initial indicator of such correlations.

Table 5. Multicollinearity diagnostic tests in the Turkish Women's Volleyball League (with observations 145 and 286 removed)

Variable	VIF	SQRT VIF	Tolerance	R-Squared			Eigenval	Cond index
LPE	1.09	1.04	0.9179	0.0821		1	3.0789	1.0000
SE	1.01	1.01	0.9863	0.0137		2	0.9734	1.7785
MBE	1.06	1.03	0.9400	0.0600		3	0.6709	2.1423
UPE	1.01	1.01	0.9889	0.0111		4	0.6119	2.2432
OHE	1.12	1.06	0.8898	0.1102		5	0.4758	2.5439
Mean VIF	1.06					6	0.1892	4.0336
						Condition number		4.0336

As shown in Table 5, the VIFs are very close to 1, the tolerances are approximately or slightly less than 1, the R^2 values are small, all eigenvalues except the first are small and the conditional indexes are below 10. All these results confirm that there is no multicollinearity between the independent variables, and the prediction model can be considered as successful.

The predicted efficiency model successfully passed all diagnostic tests. Therefore, we can reliably state that *result* is a function of LE, PE, OOE, PCE and N4E. Next, the extent to which the predicted efficacy model can accurately predict the match result was examined in a classification table. The classification table for the TWVL EM is given as Table 6.

Table 6. Classification table for the efficacy model in the Turkish Women's Volleyball League

	True		
Classified	D	~D	
+	125	27	152
–	22	122	144
Total	147	149	296
Classifield+if predicted Pr (D) > = 0.5			
True D defined as sonuç ! = 0			
			%
Sensitivity	Pr (+ I D)		85.03
Specificity	Pr (– I ~D)		81.88
Positive predictive value	Pr (D I +)		82.24
Negative predicttive value	Pr (~D I –)		84.72
False – rate for true ~D	Pr (+ I ~D)		18.12
False – rate for true D	Pr (– I D)		14.97
False + rate for classified+	Pr (~D I +)		17.76
False – rate for classified -	Pr (D I –)		15.28
Correctly Classified			83.45

According to these findings, the logistic regression model correctly classified 83.45% of the games won and lost in the TWVL. The sensitivity and specificity (proportion of lost games classified as lost) of the model were 85.03 and 81.88%, respectively.

The same analyses were then performed for the TMVL. The results of the logistic progression model for TMVL are presented in Table 7.

To identify the anomalous observations, we calculated the standardised Pearson residuals and plotted them against the observations. Some of the game results stray far from the main group of observations. For these observations, the model predicted a lost match as won with a very high probability or vice versa. Table 7 has the power to identify all prediction model's variables on the 0.05 significance level. On the other hand, with 1 point increase at efficiency values, all variables are constant without Odds Ratio values which increased by 1.290 at LPE, 1.125 at SE, 1.082 at MBE, 1.156 at OHE and 1.152 at UPE on the bet levels, between match results as won rather than loss.

Table 7. Logistic regression prediction model for the Turkish Men's Volleyball League

			Number of obs	=	296	
			LR chi^2	=	136.49	
			Prob > chi^2	=	0.0000	
Log likelihood = −136.92593			Pseudo R^2	=	0.3326	
Result	Odds ratio	Std. err.	Z	P > |z|	[95% Conf. interval]	
LPE	1.2903700	0.1259308	2.61	0.009	1.0657200	1.5623750
SE	1.1256910	0.0555178	2.40	0.016	1.0219720	1.2399360
MBE	1.0828660	0.0275310	3.13	0.000	1.0302290	1.1381930
OHE	1.1569310	0.0244180	6.91	0.000	1.1100490	1.2057940
UPE	1.1526960	0.0307475	5.33	0.000	1.0939800	1.2145630
_cons	0.2838023	0.0936025	−3.82	0.000	0.1486881	0.5416963

Table 8. Anomalous observation analysis in the Turkish Men's Volleyball League

Obs	stdres	p	Result	LPE	SE	MBE	OHE	UPE
773	−35444010	0.9255406	0	−1	1	9	20	2
895	−3.058787	0.9027569	0	−1	2	6	12	9
711	−3.039091	0.9012191	0	0	3	1	14	7
722	−2.836845	0.8886646	0	−1	0	9	9	11
933	−2690464	0.8763931	0	−2	−1	9	1	21
802	2.118333	0.1887364	1	0	−2	−9	12	−7
770	2.199616	0.1729381	1	−1	−2	−1	−5	7
917	2.504905	0.1397109	1	0	−4	8	−3	−2
935	3.898944	0.0629296	1	−6	−6	5	−7	10
901	4.069080	0.0576677	1	−3	−3	11	−3	−6

As shown in the anomalous observation analysis (Table 8), the teams in observations 773, 895, 711, 722 and 933 lost the matches they were expected to win; conversely, the teams in observations 802, 770, 917, 93,5 and 901 won the matches they were expected to lose. After analysing these anomalous observations, observations 901 and 935 passed the criteria for influential observations in three different checks, as seen in Table 9.

Table 9. Analysis of efficient observations in the Turkish Men's Volleyball League

Obs	pgbon	stdres	dv
901	0.2195599	4.069080	2.388748
935	0.3174982	3.898944	2.351909

As shown in the table 9; 901 and 935 observations are the most efficient observations which could affect the prediction model.

Table 10 presents the results of the logit regression model after removal of the efficient observations 901 and 935, and Table 11 summarises the results of the multicollinearity analysis. The classification table for the TMVL EM is given as Table 12.

Table 10. Results of logit regression model in the Turkish Men's Volleyball League after removing the efficient observations (observations 901 and 935)

				Number of obs	=	294
				LR chi^2	=	145.79
				Prob > chi^2	=	0.0000
Log likelihood = −130.88419				Pseudo R^2	=	0.3577
Result	Odds ratio	Std. err.	Z	P > |z|	[95% Conf. interval]	
LPE	1.3761730	0.1427224	3.08	0.002	1.12304	1686362
SE	1.1475500	0.0585278	2.70	0.007	1.038385	1.268192
MBE	1.0782460	0.0279459	2.91	0.004	1.024841	1.134434
OHE	1.1690990	0.0259125	7.05	0.000	1.119399	1.221006
UPE	1.1641720	0.0323627	5.47	0.000	1.10244	1.229362
_cons	0.2784306	0.0944237	−3.77	0.000	0.1432354	0.541232

It seems that the results of logit regression prediction results in the table 7 with 296 compared observed results, without 901 and 935 observations, model's adaptability has improved. The value of LR chi^2 = 136.49 has been raised to LR chi^2 = 145.79; the value of Pseudo R^2 = 0.3326 has been raised to Pseudo R^2 = 0.3577. Excluding these two observations, effects logit regression factors as impoving explanatory power of the LPE's result variable from p> |z| = 0.009 to p > |z| = 0.002. At the same time all variables in this table has the power to identify efficiency model. Also Odds Ratio bet results has increased at all variables without LPE.

Table 11. Multicollinearity diagnostic tests in the Turkish Men's Volleyball League (after removing observations 901 and 935)

Variable	VIF	SQRT VIF	Tolerance	R-Squared		Eigenval		Cond index
LPE	1.01	1.00	0.9922	0.0078		1	2.8614	1.0000
SE	1.10	1.05	0.9104	0.0896		2	1.0694	1.6358
MBE	1.04	1.02	0.9657	0.0343		3	0.7252	1.9864
OHE	1.05	1.02	0.9558	0.0442		4	0.6708	2.3797
UPE	1.04	1.02	0.9649	0.0351		5	0.5053	2.3797
Mean VIF	1.04					6	0.1679	4.1281
						Condition number		4.1281

It has confirmed that there is no multiple linearity between the model's arguments with Table 11; VIF (Variance Sweeling Factors) are close to 1, tolarance values are below 1 and close to 1, R^2 values are minor, eigenvalues (Eigenval) are majör at the first value and the condition indexes (Con Index) are smaller than 10.

Table 12. Classification table for the efficacy model in the Turkish Men's Volleyball League

	True		
Classified	D	~D	
+	115	33	148
−	31	115	146
Total	146	148	294
Classifield+if predicted Pr (D) > = 0.5			
True D defined as sonuç ! = 0			
			%
Sensitivity	Pr (+ I D)		78.77
Specificity	Pr (− I ~D)		77.70
Positive predictive value	Pr (D I +)		77.70
Negative predicttive value	Pr (~D I −)		78.77
False − rate for true ~D	Pr (+ I ~D)		22.30
False − rate for true D	Pr (− I D)		21.23
False + rate for classified +	Pr (~D I +)		22.30
False − rate for classified −	Pr (D I −)		21.23
Correctly Classified			78.23

The model correctly classified 78.23% of the wins and losses in TMVL. The sensitivity and specificity of the model in TMVL were 78.77 and 77.70%, respectively.

Discussion

There are several researches in the literature to predict the match result in volleyball, but the number that evaluates them in a model that includes many parameters are very less.

Nikos, Elissavet (2011) investigated the relationship between impulse performance and attack tempo and they found that the success rate of the offenses with quick tempo attackers was more high and setter's performance and attack tempo as determinants of attack efficacy in Olympic level male volleyball teams.

In a similar manner, Marcelino, Sampaio, Mesquita (2012) investigated the relationship between attack and serve performance in different periods of time and opponent team level from 600 rallies which were selectively sampled from a total of 5.117 rallies at Men's World Cup 2007 and determined that service and attack performance changed in relation to both opponent team level and match period.

Silva, Lacerda, Joao (2013) analyzed 24 matches during the men's Senior Volleyball World Championship-Italy 2010 for to understand what happens when the setter is in the attack zone (zones 4, 3 and 2). They determined that the team had a negative effect on the successful outcome when setter came to attack zone position (zones 4, 3, 2) because the team could not effectively perform the offensive skills needed to win such as offensive side out.

In the Men's Volleyball World Championship, the outcomes were best explained by attack error, jump serve point, quick ball error and jump serves (Asterios et al., 2009). Analysing the Greek premier league male volleyball teams during 2005–2006, Drikos et al. (2009) found that team performance was most affected by the serve efficacy ratio and attack efficacy ratio.

In female volleyball, Monteiro et al. (2009) linked the set outcome to attack efficacy. Marelic et al. (2004) concluded that all their tested technical elements (serve, serve reception, spike in the phase of attack, block and spike in the phase of counter attack variables) effectively determine the game outcomes.

Pena, Rodriguez-Guerra, Busca, Serra (2013) studied on high-level men's volleyball in the Spanish Superliga during the 2010–2011 season, to see which skills and factors were more effective to predict winning and losing. They found that the points obtained in the break point phase, number of reception errors and the number of blocked attacks by the opponent were significant predictors of winning or losing the matches.

Gonzalez-Silva, Domingues, Fernandez-Echeverria, Rabaz, Arroyo (2016) sampled 5,842 game actions carried out by the 16 male category and the 18 female category teams that participated in the Under-16 Spanish Championship for to evaluate setting efficacy in young male and female volleyball players. Results showed that the best predictive variables of setting efficacy, both in female and male categories, were reception efficacy, setting technique and tempo of a set.

Rentero, Joao, Moreno (2015) analyzed libero's participation and their influence in the attack and defense phases in men's elite volleyball with the sample of 1,101 pass and defense game actions of the four highest-placed teams in the 2008 Beijing Olympic Games. The results of the study revealed that there were significant associations during the defensive stage of the game between the defending player and the defensive phase, the libero's defence predominating in zone 5; the defending player and defense efficiency, which is improved by the libero; the defending player and counterattack, as attacks increased in zone 6 when the libero was defending.

In another study of Silva, Lacerda, Joao (2014) shows that service points, reception errors, and blocking errors were the discriminating variables that identify the final outcome of the match (won/loss). Moreover, successful service points were the major variable most likely associated with match success (won). In this sense, increasing the effectiveness of service should be a top priority in coaching elite volleyball teams.

As seen in the literature examples above, the results of the match or the factors that determine whether to win or loss are always influenced by the researchers. In the researches conducted, one or more volleyball techniques or the effects of the player's positions were searched and the results were obtained. This study differs from others in that it contains all the techniques and player's positions in volleyball.

To test the success of the predicted efficacy model in predicting the match result, we compared the predicted and actual results. The model correctly predicted the outcomes of 83.45% of the games in TWVL and 78.23% of the games in TMVL. In other words, the probability of correct prediction by the model is 83.45% in women's league and 78.23% in men's league.

The results of this study were derived from the match data of teams in the Turkish Elite Volleyball League. However, to determine whether these results characterise the league, the same league should be reanalysed for both genders using data of the 2014–2015 season. Repeating the study for the elite men's volleyball league would determine whether the results are generalisable and would highlight the gender differences. These analyses have been planned as extensions to this study and are currently being investigated.

Conclusion

Statistical reports of sports matches contain many variables. Such numerous and varied data on the teams and individual players are difficult to organise into a performance evaluation by match analysis programs. Therefore,

by identifying the variables that significantly establish the match result, coaches can develop the contents of both development and tactical training programs.

In this context, Akarçeşme (2010) determined that the match results were well explained by an efficacy model based on the team's position and the technical skills that determine success or failure. The validity and reliability of the model was confirmed in an analysis of the next two years' worth of league data. The model can also explain gender differences in volleyball matches.

Recommendations

The data in the present study were sampled from Turkish men's and women's volleyball leagues in 2010 and 2011. In that season, teams in both leagues were entitled to align 3 foreign players and the teams were found to exercise this right. The study could be repeated for seasons in which teams were entitled to align 2 foreign players. Currently, we are analysing the data of the 2011–2012 and 2012–2013 seasons. These additional analyses will further validate the model and will reveal the tendencies of the leagues employing various numbers of foreign players, in terms of the EM variables.

Acknowledgement

References

Akarçeşme, C. (2010). *Determination of the variables explaining the match result with the logistic regression method and estimation of the probability model aimed at winning the match in the elite women's volleyball*. Unpublished Doctorate Thesis. Ankara: Gazi University Institute of Health Sciences.

Arauj, R., Mesquit, I., Marcelino, R. (2009). Relationship between Block Constraints and set outcome in elite male volleyball. *International Journal of Performance Analysis in Sport*, *9* (8), 306–313.

Asterios, P., Kostantinos, C., Athanasios, M., Dimitrios, K. (2009). Comparison of technical skills effectiveness of men's national volleyball teams. *International Journal of Performance Analysis in Sport*, *9*, 1–7.

Beim, G., Winter, R. (2003). *The female athlete's body book: How to prevent and treat sports injuries in women and girls*. New York: The McGraw Hill Companies.

Drikos, S., Kountouris, P., Laios, A., Laios, Y. (2009). Correlates of team performance in volleyball. *International Journal of Performance Analysis in Sport*, *9*, 149–156.

Gonzalez-Silva, J., Dominguez, A.M., Fernandez-Echeverria, C., Rabaz, F.C. Arroyo, M.P.M. (2016). Analysis of setting efficacy in young male and female volleyball players. *Journal of Human Kinetics, 53* (1), 189–200.

Gualdi-Russo, E., Graziani, I. (1993). Anthropometric somatotype of Italian sport participants. *The Journal of Sports Medicine and Physical Fitness*, *33*, 282–291.

Hosmer, D.W., Lemeshow, S. (2000). *Applied Logistic Regression* (2 ed.). New York: Wiley-Interscience Publication.

Korhonen, M.T., Mero, A., Suominen, H. (2003). Age-related differences in 100-m sprint performance in male and female master runners. *Medicine and Science in Sports and Exercise*, *35*, 1419–1428.

Marcelino, R.O., Sampaio, J.E., d Mesquita, I.M. (2012). Attack and serve performances according to the match period and quality of opposition in elite volleyball matches. *Journal of Strength and Conditioning Research*, *26* (12), 3385–3391.

Marelic, N., Resetar, T., Jankovic, V. (2004). Discriminant analysis of the sets won and the sets lost by one team in A1 Italian volleyball league – A case study, *Kinesiology, 36*, 75–82.

Monteiro, R., Mesquita, I., Marcelino, R. (2009). Relationship between the set outcome and the dig and attack efficacy in elite male volleyball game. *International Journal of Performance Analysis of Sport*, *9*, 294–305.

Nikos, B., Elissavet, N.M. (2011). Setter's performance and attack tempo as determinants of attack efficacy in Olympic-level male volleyball teams. *International Journal of Performance Analysis in Sport*, *11* (3), 535–544.

Pena, J., Rodriguez-Guerra, J., Busca, B., Serra, N. (2013). Which skills and factors better predict winning and losing in high-level men's volleyball? *Journal of Strength and Conditioning Research*, *27* (9), 2487–2493.

Pregibon, D. (1981). Logistic regression diagnostic. *The Annals of Statistics*, *9*, 705–724.

Rentero, L., Joao, P.V., Moreno, M.P. (2015). Analysis of the libero's influence of the libero in different phases of game in volleyball. *Revista Internacional De Medicina Y Ciencias De La Actividad Fisica Y Del Deporte, 15* (60), 739–756.

Rickenlund, A., Carlstrom, K., Jörn Ekblom, B., Brismar, T.B., von Schoultz, B., Hirschberg, A.L. (2003). Hyperandrogenicity is an alternative mechanism underlying oligomenorrhea or amenorrhea in female athletes and may improve physical performance. *Fertility and Sterility*, *79*, 947–955.

Silva, M., Lacerda, D., Joao, P.V. (2013). Match analysis of discrimination skills according to the setter attackzone position in high level volleyball. *International Journal of Performance Analysis in Sport, 13* (2), 452–460.

Silva, M., Lacerda, D., Joao, P.V. (2014). Game-related volleyball skills that influence victory. *Journal of Human Kinetics*, *41* (1), 173–179.

Weisberg, S. (2005). *Applied Linear Regression* (3 ed.). New Jersey: Wiley & Sons Inc. Publication.