OPERATIONS RESEARCH AND DECISIONS

No. 4 2016
DOI: 10.5277/ord160402

Ananda BANDULASIRI[1]
Tom BROWN[2]
Indika WICKRAMASINGHE[3]

FACTORS AFFECTING THE RESULT OF MATCHES IN THE ONE DAY FORMAT OF CRICKET

Factors contributing to winning games are imperative, as the ultimate objective in a game is victory. The aim of this study was to identify the factors that characterize the game of cricket, and to investigate the factors that truly influence the result of a game using the data collected from the Champions Trophy cricket tournament. According to the results, this cricket tournament can be characterized using the factors of batting, bowling, and decision-making. Further investigation suggests that the rank of the team and the number of runs they score have the most significant influence on the result of games. As far as the effectiveness of assigning bowlers is concerned, the Australian team has done a fabulous job compared to the rest of the teams.

Keywords: *victory in a cricket match, efficient bowling, assignment problem, team sport*

1. Introduction

Cricket is a bat and ball team-game with three major formats: test cricket, ODI cricket, and twenty20 cricket. A test match is played for five days, A one Day International (ODI) match is played for 50 overs[4] per side, and a twenty20 is played for 20 overs per side. Each team consists of 11 players and the team which bats first is decided by tossing a coin. In ODI matches, the team batting first continues scoring runs

[1]Department of Mathematics and Statistics, Sam Houston State University, Huntsville, Texas 77341, USA, e-mail address: abandulasiri@shsu.edu

[2]Department of Mathematical Sciences, Eastern New Mexico University, Station 18, 1500 S Ave K, Portales, NM 88130, USA, e-mail address: tombrown@enmu.edu

[3]Department of Mathematics, Prairie View A&M University, P. O. Box 519-MS 2225, Prairie View, TX 77446-0519, USA, e-mail address: wickramasingheindika@gmail.com

[4]Six deliveries bowled by the bowler (pitcher).

until they have used up the permitted overs or they lose 10 wickets. In order to win the match, the team batting second needs to exceed the number of runs scored by the first team within the 50 overs without losing 10 wickets. ODI was the most entertaining format, due to the dynamic nature of the game, before the introduction of twenty20. With the continued popularity of the ODI format, there is a good demand for studies that focus on the performances of both the players and teams.

Though the game of cricket mainly depends on the three departments of batting, bowling, and fielding, on a given day a mixture of unforeseen factors can influence the result. In addition, the victory in a match highly depends on accurate decision taking. There is a series of decisions a team is required to take, including picking the correct combination of players for the match, reading the behavior of the turf[5] precisely, organizing the batting and fielding positions, selecting the appropriate bowler to bowl decisive overs in the match, and guiding the team to the victory. Due to the multi-dimensional nature of decision making in cricket, characterizing a game and identifying the factors that influence the result of a match is complex but beneficial for the game. Though there are several benefits of such studies, two of them are as follows: First, characterizing a game is helpful for both the captain and coach of a team to devise their game plan. Secondly, if proper characterization of the game can be done, the result of a match can be predicted with a higher accuracy.

The International Cricket Council (ICC), the game's governing body, was the organizer of the 2013 Champions Trophy. This was held in England during June with the participation of eight teams, namely Australia (AUS), England (ENG), India (IND), New Zealand (NZ), Pakistan (PAK), Sri Lanka (SL), South Africa (SA), and the West Indies (WI). India, England, South Africa, and Sri Lanka qualified to play the semi-final matches. Thereafter, both the Indian and English teams successfully progressed to the final, with the Indian team winning the 2013 Champions Trophy.

With the popularity of the game, studies that aim to predict the result of an ODI match have become very important. According to the literature, some of the most frequently used variables to predict the result of ODI games are home-field advantage [1, 2, 8, 13, 17], result of the coin toss [5, 6, 8, 15], day/night effect [6], the effect of bowling [11] and batting [3, 9, 10, 11, 12, 19, 20]. It has become common practice to use the above mentioned factors in most of the studies related to cricket. The literature indicates that minimal attention has been given to the impact of other indicators such as the rank of the opposing team, batting partnerships, assignment of bowlers by the team's captain, and the composition of teams towards the result of a match.

The rest of the paper is organized as follows. Section 2 presents some descriptive statistics about the collected data set. Section 3 discusses the characterization of the game, section 4 compares the relevance of batting partnerships, and section 5 conducts

[5] The surface that the match is played on.

an in-depth study about the assignment of bowlers by the team captain. Finally, section 6 concludes the study.

2. Collected data

In this study, the 30 ODI innings that took place in the 2013 Champions' Trophy tournament were considered for the investigation. The number of fifties (at least 50 runs) made by batsmen in the match (Fifties), number of partnerships (Partnership) over fifty runs, number of spinners[6] (NoofSpinners) played in the match, number of fast bowlers played (NoofFastBowlers), and the number of all-round players (NoofAllRounders) played are some of the considered characteristics. In this study, a player who bats at or before the 7th position in the batting order is classified as a batsmen. In addition, a player who represented the team as a batsman and bowled at least 3 overs was considered to be an all-rounder for the analysis. The result of each match was recorded as 0 (a non-win) and 1 (a win). The ranks of each of the eight teams were recorded on the basis of ICC publications.

Table 1. Descriptive statistics

Variable	Mean Winning	Mean Non-Winning	Significance
Fifties	1.47	1.00	0.24
Partnership	1.53	0.86	0.02
NoofSpinners	2.07	1.71	0.25
NoofFastBowlers	3.27	3.71	0.06
NoofAllRounders	2.01	2.43	0.25

Table 1 summarizes the average values of the selected variables for both winning and non-winning teams. The average number of fifties scored in an innings by winning teams (1.47) is higher than that of non-winning teams (1.00), though the difference is not statistically significant. The difference between the number of partnerships scoring more than 50 runs for winning and non-winning teams is statistically significant ($p < 0.05$). The average number of spinners bowling in an innings for winning teams was 2.07, while this value for non-winning teams was 1.71. Hence, unsurprisingly, the number of fast bowlers used in winning teams (3.27) is lower than in non-winning teams (3.71). Non-winning teams used more all-rounders[7] (2.43) on average than their counter parts (2.01). As Table 1 indicates, there are differences in the number of spinners used,

[6]A bowler who bowls the cricket ball with rapid rotation, thus gaining sideways movement after the ball hits the turf.

[7]A cricketer who is good at all the departments of the game.

the number of fast bowlers, and the number of all-rounders for winning and non-winning teams, though they are not statistically significant.

3. Characterization of an ODI matchz in the Champions' Trophy

In order to characterize a game in this selected tournament, the Principle Component Analysis (PCA) was utilized. Since data on a large number of performance variables are available for the game of cricket, PCA enables us to select the most related variables to characterize the game. One of the reasons for the popularity of this technique is due to its numerical feasibility for reducing higher dimensional systems [4, 18, 22]. A glance at the literature reveals the application of PCA in sports related studies. Wittea et al. [22] discussed the application of PCA in the context of movement coordination in sport. In another study, Manage et al. [14] demonstrated a simple and straightforward technique, which is based on PCA to quantify a player's performance in the 2012 T20 Cricket World Cup. In addition to these, Nalik and Khattree [16] illustrated how to use PCA with sports data.

The main purpose of the application of PCA is to reduce the number of variables. This technique converts a large number of possibly correlated variables into a set of linearly uncorrelated variables in an optimal way. These new variables, of which there are fewer than in the original set, are called principal components. This is achieved by a series of vector transformations at the end of which; the higher dimension is reduced to a smaller dimension, so that the variability of the data is explained by a smaller number of variables.

Consider a sample of n observations of a vector of $p = 6$ variables, $X = (X_1, X_2, \ldots, X_p)$.

The first principal component is calculated according to the following steps. Here X_1 = BatFirst, X_2 = NoofSpins, X_3 = NoofFst, X_4 = Fifties, X_5 = NoofAllround, and X_6 = Partnerships.

- Start with the variables $X = (X_1, X_2, \ldots, X_p)'$.
- Find the rotations of these variables $Y_1, Y_2, ..., Y_p$ called principal components, which satisfy the following two conditions:

i) $Y_1, Y_2, ..., Y_p$ are uncorrelated;

ii) $\mathrm{var}(Y_1) \geq \mathrm{var}(Y_2) \geq ... \geq \mathrm{var}(Y_p)$.

- When calculating Y_j, we need to find $a'_{(j)} = (a_{1j}, a_{2j}, ..., a_{pj})$ such that $Y_j = a'_{(j)}X$ and $\sum a_i^2 = 1$.

• When calculating Y_{j+1} we need to find $a'_{(j+1)} = (a_{1(j+1)}, a_{2(j+1)}, ..., a_{p(j+1)})$ such that $Y_{j+1} = a'_{(j+1)} X$.

• Y_j and Y_{j+1} are uncorrelated.

Table 2. Total variance explained

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	1.825	30.415	30.415	1.825	30.415	30.415	1.659	27.642	27.642
2	1.634	27.241	57.656	1.634	27.241	57.656	1.472	24.541	52.183
3	1.136	18.938	76.593	1.136	18.938	76.593	1.465	24.410	76.593
4	0.680	11.338	87.931						
5	0.435	7.254	95.186						
6	0.289	4.814	100.000						

According to Table 2, there are three components with eigenvalues greater than 1. This indicates the existence of three factors: factor 1, factor 2, and factor 3, which explain 30.42, 27.24, and 18.94% of the total variation, respectively. Factor 1 accounts for more variance than the other two factors. The cumulative variance accounted for by the three factors is 76.59%. Based on the high factor loadings displayed in Table 3, factor 1 is influenced by the variables NoofAllround, Fifties, and Partnership. As these variables are directly related to batting, factor 1 is identified as Batting-Performance. Variables related to factor 2 are NoofSpins and NoofFst. It is obvious that these variables reflect the bowling strength of a team. Therefore, the second factor is identified as Bowling-Performance. Finally, factor 3 comprises the following variables: BatFirst, NoofSpins and NoofAllround. It is a well known fact that English cricket grounds are more helpful to fast bowlers than spinners. Therefore, playing spinners in a side is not common on English turfs. When a team is selected based on the opposing team, spinners are selected for tactical reasons. Though it is clear that batting first depends on the outcome of the toss, even after winning the toss the captain has a huge responsibility to take the correct decision on this matter. In addition, selecting the appropriate number of all-rounders in a team is done as a part of the game plan. It is evident that all these variables are directly or indirectly related to the decision making aspect of the game. Hence, factor 3 is named Decision-Making. Therefore, an ODI game in this tournament can be characterized by batting, bowling, and decision-making.

In this study, PCA was conducted on the above selected items with orthogonal rotation. According to the Kaiser–Meyer–Olkin measure, the sampling adequacy is given by KMO = 0.553. As Field [7] states, this value of 0.553 is sufficiently large. Bartlett's Test of Sphericity, $\chi^2(21) = 26.365$, $p < 0.05$, indicates that the correlation of items is sufficiently large. The three factors retained in the final analysis and the factor loadings after the rotation are presented in Table 3.

Table 3. Summary of explanatory factor analysis

Item	Rotated factor loadings		
	Factor 1 Batting-performance	Factor 2 Bowling-performance	Factor 3 Decision-making
BatFirst	–0.009	0.312	–0.775
NoofSpins	–0.224	0.626	0.634
NoofFst	0.444	–0.715	–0.083
NoofAllround	0.714	–0.166	0.593
Fifties	0.781	0.353	0.020
Partnership	0.727	0.346	–0.072

After characterizing the game, it is important to identify the variables that truly impact the likelihood of winning a game. Therefore, after running logistic regression, only Partnership ($\beta = 2.35$, $p < 0.05$) significantly contributes to the model. While the contribution of NoofFst, and NoofAllround is negative, Partnership and Fifties positively contribute to the probability of winning. When the number of partnerships scoring over 50 runs is increased by one, the odds of winning a match are increased by 7.85 times given the other variables are kept constant.

4. Number of runs scored by batting partnerships and the rank of the team

In the game of cricket, two batsmen always bat together in a partnership until the partnership breaks. This partnership between two batsmen comes to an end after one of the batsmen gets out (in cricketing terms a wicket falls) and is replaced by a player who has not yet batted, thus forming a new pair. As there are 11 batsmen in a team, at most 10 partnerships can be formed in the course of an innings. In cricketing terms, the number of runs scored by the *i*th pair is referred to as the *i*th wicket partnership. In this tournament, as the statistics indicate, the average number of wickets of the winning team that fell in a match is approximately 6. Therefore, a comparison of the distribution of partnerships is made for the first six wickets. Figure 1 shows the distributions of these partnerships. As Tan and Zhang [20] stated, the first wicket partnership is probably one of the most important ones. According to Fig. 1A, the mean first wicket partnership of teams who have lost matches is lower than that for winning teams. In particular, it shows that the probability of partnerships of over 50 runs is higher for winning teams than that for non-winning teams. For the second wicket (Fig. 1B), as expected, winning teams have performed better than non-winning teams, especially with regards to the probability of second wicket partnerships of over 60 runs.

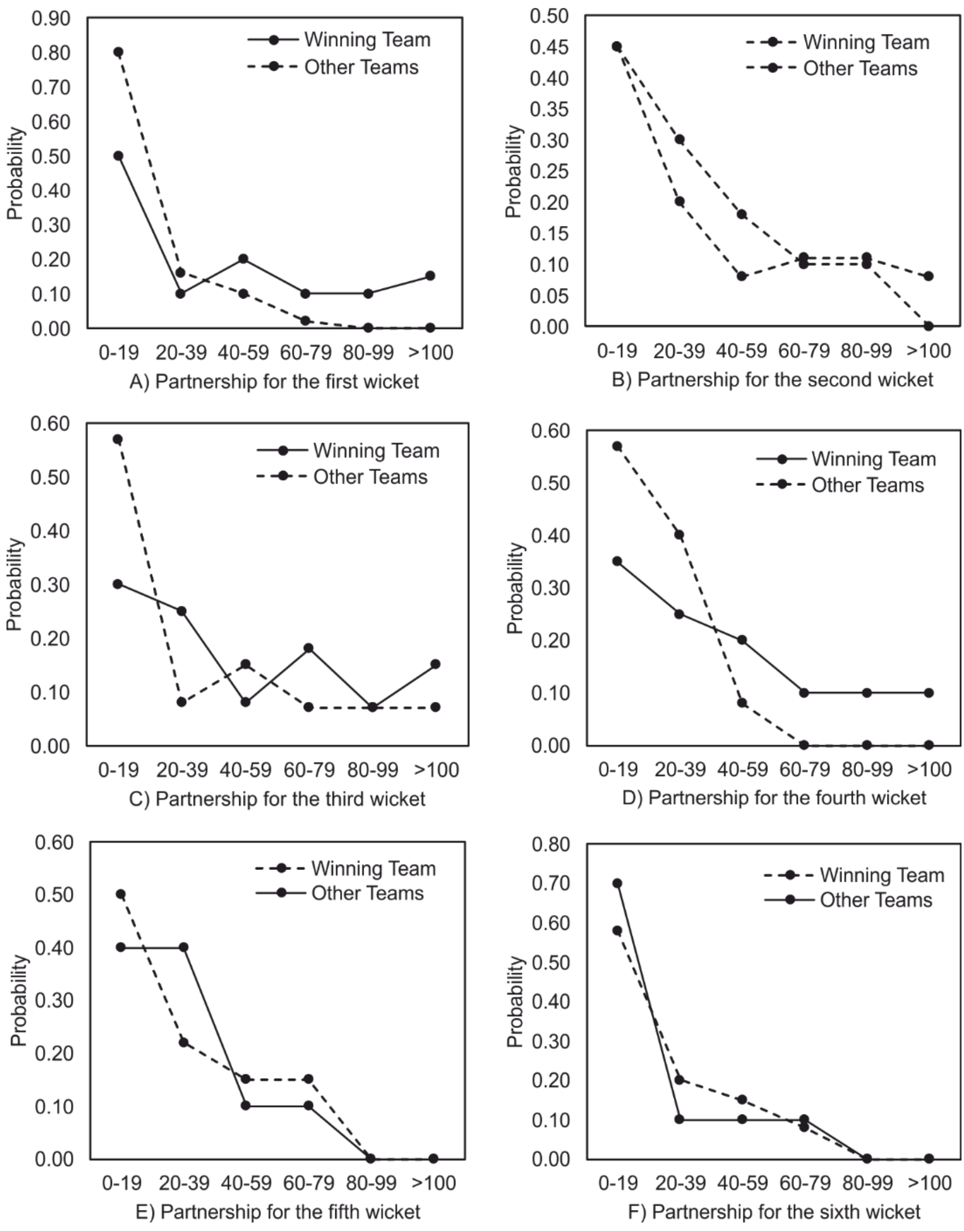

Fig. 1. Probability Distributions of Partnerships (solid line-winning team, dashed line- non-winning team)

The behavior of the distribution of partnerships for the third wicket shows an irregular pattern in comparison with the rest. Nevertheless, Fig. 1C shows that winning teams have a better record of partnerships of over 60 runs than their opposing teams. Figure 1D shows the distribution of partnerships for the fourth wicket. According to this, teams who have lost matches have not had any partnerships of over 60 runs, but their opposing teams have had better partnerships.

In an ODI cricket match, the middle-order partnerships are crucial for winning a match. Most importantly, teams who bat in the second innings need to control the pace of the game until they achieve the target. The probability distributions of partnerships for the fifth and sixth wickets are different from the rest of the distributions. A comparison of these distributions indicates very little difference between winning and losing teams in the number of runs scored by these partnerships.

Combining these results, it can be inferred that the first four partnerships are important for setting a base for the innings as a whole. As the number of overs progresses, if a team has a large number of wickets remaining, then the batsmen can implement a risky strategy, i.e. trying to score runs quickly, while accepting a higher probability of getting out. Hence, later on in the innings the rate at which runs are scored and not the number of runs scored by a partnership becomes more important.

When considering the effect of a team's rank on the result of a game, data show that there is a significant and positive correlation between the number of matches won and the rank of the team ($\gamma = 0.717$, $p < 0.05$). According to the collected data, teams with a ranking of over 110 (high rank) recorded more wins than losses, while teams with a ranking close to 100 show more losses than wins. This is evident when considering teams such as Pakistan (94) and South Africa (98). As expected, teams with a lower ranking (less than 80) recorded more losses than wins.

5. Effectiveness of assigning the bowlers to bowl 50 overs in a game

As everyone knows, bowling is an integral part of cricket and so it is in this tournament. Bowling performance in the Champions' Trophy has been extensively analyzed by Wickramasinghe [21]. Therefore, we do not repeat a similar investigation here but explore one of the most important aspects that most of the prior studies have not considered, namely the effectiveness of the assignment of bowlers by the team's captain. For this investigation, the five bowlers from each team who bowled the largest number of overs in each of the 10-over slot (1st–10th, 11th–20th, 21st–30th, 31st–40th and 41st–50th) were selected. The allocation of a bowler to bowl in the ith 10-over slot can be seen as an assignment problem in operations research. In this case, there are five workers (bowlers) and five jobs (to bowl the largest number of overs in the ith 10-over slot). Then it is necessary to find the best combination of bowlers to get the job done, so that this assignment attains the least possible cost. In this situation, the Hungarian method, which is discussed below, is appropriate to use.

Let C_{ij} represent the cost of assigning the ith bowler to bowl the jth 10-over slot and define

$$x_{ij} = \begin{cases} 1, & \text{if the } i\text{th bowler bowls the } j\text{th 10-over slot} \\ 0, & \text{otherwise} \end{cases}$$

Thus the assignment can be represented as a standard Linear Programing (LP) problem as follows.

$$\text{Minimize the total cost } Z = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij}$$

$$\text{subject to } \sum_{j=1}^{n} x_{ij} = 1, \; i = 1, 2, ..., n, \; x_{ij} = 0 \text{ or } 1 \tag{1}$$

Definition of the cost matrix C

When quantifying a bowler's performance, the literature shows the use of several commonly used statistics such as the average number of runs per over, the number of wickets, and strike rate (the average number of balls bowled per wicket). As Lemmer and Lewis [11, 12] argued that the above traditional performance indicators for bowlers are not appropriate to give a single overall measure of bowling performances, the CBR^* statistic, introduced by Lemmer [11], is used as a means of quantifying a bowler's performance. A better bowler is expected to have a smaller CBR^* value. When preparing the cost matrix **C** the five bowlers from each team who bowled the most overs in each 10-over slot were selected. Using the statistics for these bowlers', the CBR^* values were calculated according to the following formula.

$$CBR^* = \frac{3R}{W^* + O + \dfrac{W^* R}{B}} \tag{2}$$

where O denotes the number of overs the bowler bowled and W^* is the sum of the weights [11] for each wicket taken by the bowler. In addition, R and B represent the number of runs conceded by the bowler and the number of balls bowled by the bowler. The value C_{ij} in the cost matrix **C** is given by CBR^*_{ij} value which is the CBR^* value for the ith bowler when bowling in the jth 10-over slot, and a C_{ij} matrix was prepared for each team. For future reference, the total cost of assigning 5 bowlers by each team is called the cost of bowling assignment (CBA).

Table 4. Assignment of bowling

Country	CBA
Australia	27.597
Sri Lanka	29.778
West Indies	31.958
England	36.100
Pakistan	40.424
India	42.956
New Zealand	43.553
South Africa	52.082

The CBA value is smallest (27.597) for the Australian team while it is largest (52.08) for the South African team. According to these statistics, the Australian team was the best able to successfully assign their bowlers in the tournament. Though the South African team is renowned as having one of the best bowling units in the world, they did not perform well in this tournament. All the statistical calculations were conducted using R Studio and Minitab.

6. Conclusion

Characterizing success statistics for the ODI format of cricket is important in many ways. This study used statistics from the 2013 Champions Trophy to undertake this analysis. According to PCA, the cricket matches played in this tournament can be characterized as a combination of three factors. These three factors were identified as batting, bowling and decision-making, and they accounted for approximately 30, 27, and 19% of the total variation, respectively. Out of the various aspects considered, there is strong evidence that only the number of partnerships of over 50 runs and the rank of the opposing team have an influence on the result of a match in this tournament. When further attention is given to partnerships, as expected the winning team has been able to maintain better partnerships for most of the initial wickets.

Bowling is a key discipline of the game of cricket. One of the critical, but rarely investigated, aspects of the game of cricket is the assignment of bowlers by the team captain. According to the results given by the Hungarian method, using CBR^* values, the Australian team did a fabulous job in assigning bowlers. A reasonable question appears as to why the Australian team did not qualify to play in the semi-finals, though they were the best team as far as the bowling allocation is concerned. The simple answer is their poor batting performance in the tournament. When the top ten most successful batsmen in this tournament are considered, none of the Australian players can be seen in that list. Both Sri Lanka and England played in the semi-finals and they have the 2nd and 4th lowest CBA values (approximately 30 and 36 respectively). Although the Indian team won the tournament, they did not have a very good bowling allocation according to the CBA values. As opposed to the Australian team, the Indian batsmen performed really well in the tournament. There are three Indian batsmen (Dhawan, Sharma, and Kohli) who were among the top five scorers of the tournament. Therefore, it is clear that success in this tournament did not solely depend on the bowling performances (small CBA values), as cricket is a game with three main disciplines: batting, bowling, and fielding.

Due to the nature of most cricket tournaments, it is difficult to collect larger collections of data for an investigation like this. Though the KMO value suggests that the sample size is sufficiently large for PCA, it worth collecting more samples of data for analysis. Regarding the analysis of the bowling assignment, it would be interesting to consider, e.g., 5-over slots, but due to the sample size this was not possible.

References

[1] BAILEY M., *Predicting Sporting Outcomes. A Statistical Approach*, PhD thesis. Swinburne University, Melbourne 2005.

[2] BAILEY M., CLEARKE S.R., *Predicting the match outcome in one-day international cricket matches, while the game is in progress*, J. Sports Sci. Med., 2006, 5, 480.

[3] BEAUDOIN D., SWARTZ T., *The best batsmen and bowlers in one-day cricket*, South Afr. Stat. J., 2003, 37, 203.

[4] BRAIDO P., ZHANG X., *Quantitative analysis of finger motion coordination in hand manipulative and gestic acts*, Human Mov. Sci., 2006, 22, 661.

[5] DAWSON P., MORLEY M., PATON D., THOMAS D., *To bat or not to bat. An examination of match outcomes in day-night limited overs cricket*, J. Oper. Res. Soc., 2009, 60, 1786.

[6] DE SLIVA B.M., SWARTZ T.B., *Winning the coin toss and the home team advantage in one-day international cricket matches*, New Zealand Stat., 1997, 32, 16.

[7] FIELD A., *Discovering Statistics Using SPSS*, 3rd Ed., Sage Publications, Inc., London 2009.

[8] KALUARACHCHI A., APARNA S.V., *A classification based tool to predict the outcome in ODI cricket*, 5th Int. Conference on Information and Automation for Sustainability (ICIAFs), Colombo, Sri Lanka, 2010.

[9] KIMBER A.C., HANSFORD A.R., *A statistical analysis of batting in cricket*, J. Royal Stat. Soc., 1993, 156, 443.

[10] KOULIS T., MUTHUKUMARANA S., BRIERCLIFFE C D., *A Bayesian stochastic model for batting performance evaluation in one-day cricket*, J. Quant. Anal. Sports, 2014, 10, 1.

[11] LEMMER H., *An analysis of players' performances in the first cricket Twenty20 World Cup Series*, South Afr. J. Res. Sport, Phys. Edu. Recr., 2008, 30 (2), 71.

[12] LEWIS A.J., *Towards fairer measures of player performance in one-day cricket*, J. Oper. Res. Soc., 2005, 56, 804.

[13] BANDULASIRI A., *Predicting the winner in one day international cricket*, J. Math. Sci. Math. Edu., 2006, 3 (1), 6.
[14] MANAGE A., SCRIANO S.M., HALLUM C.R., *Performance analysis of T20-World Cup Cricket 2012*, Sri Lankan J. Appl. Stat., 2013, 14 (1), 1.
[15] MORLEY B., THOMAS D., *An investigation of home advantage and other factors affecting outcomes in English one-day cricket matches*, J. Sports Sci., 2005, 23 (3), 261.
[16] NAIK D.N., KHATTREE R., *Revisiting Olympic track records. Some practical considerations in the principal component analysis*, Amer. Stat., 1996, 50 (2), 140.
[17] PAUL A., STEPHEN R., *Factors affecting outcomes in test match cricket*, Proc. 6th Australian Conf. on Mathematics and Computers in Sport, Bond University, Queensland 2002.
[18] SADEGHI H., ALLARD P., DUHAUME M., *Functional gait asymmetry in able-bodied subjects*. Human Mov. Sci., 1997, 16, 243.
[19] SCARF P., SHI X., AKHTAR S., *On the distribution of runs scored and batting strategy in test cricket*, J. Royal Stat. Soc. A, 2011, 174, 471.
[20] TAN A., ZHANG D., *Distribution of batting scores and opening wicket partnerships in cricket*, Math. Spectrum, 2001/2002, 34 (1), 13.
[21] WICKRAMASINGHE R.I.P., *Bowlers' performances in 2013 Champions Trophy*, Ann. Appl. Sport Sci., 2014, 2 (1), 1.
[22] WITTEA K., GANTER N., BAUMGART C., PEHAM C., *Applying a principal component analysis to movement coordination in sport*, Math. Comp. Model. Dyn. Syst., 2010, 16 (5), 477.
[23] WU J., WANG J., LIU L., *Feature extraction via KPCA for classification of gait patterns*, Human Mov. Sci., 2007, 26 (3), 393.

Received 10 July 2016
Accepted 29 December 2016