

## Research Article

# Rating Batters in Test Cricket

**Sohail Akhtar** <sup>1</sup>, **Muhammad Awais**,<sup>2</sup> **Ibrar Hussain**,<sup>2</sup> **Gauhar Ali**,<sup>3</sup> **Sadique Ahmad** <sup>3,4</sup>,  
**Muhammad Imran Khan** <sup>1</sup>, **Humera Qureshi**,<sup>1</sup> and **Mohammed ElAffendi** <sup>3</sup>

<sup>1</sup>Department of Mathematics and Statistics, The University of Haripur, Haripur, Khyber Pakhtunkhwa, Pakistan

<sup>2</sup>Department of Statistics, Government College University Lahore, Lahore, Pakistan

<sup>3</sup>EIAS, Data Science and Blockchain Laboratory, College of Computer and Information Sciences Prince Sultan University, Riyadh 11586, Saudi Arabia

<sup>4</sup>Department of Computer Sciences, Bahria University, Karachi Campus, Karachi, Pakistan

Correspondence should be addressed to Sohail Akhtar; s.akhtar@uoh.edu.pk

Received 8 June 2022; Accepted 21 July 2022; Published 31 August 2022

Academic Editor: Tahir Mehmood

Copyright © 2022 Sohail Akhtar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In test cricket, we rated batters' performance. We have proposed rating criteria as introduced by Scarf, Akhtar, and Rasool in 2014 with additional explanatory variables on the updated data set. The newly added covariates that we included in our research are the home factor and the ground influence. The same rating system is applied in the previous study. Using multinomial logistic regression, sessions from all days of a test match are modeled to determine match outcome probabilities at the end of each session. These models are based on all of the factors that can influence the outcome of a match. It is discovered that the predictors of home factor and pitch quality have a significant impact on the outcome of the test match. We used multinomial logistic regression to model data and estimate the parameters in the models. We forecasted match outcomes using these models at the end of each session and measured batters' performances by using these probabilities. This process is repeated in a test match at the end of a session, and batters' contributions to their team score are accumulated. Both teams' batters are then ranked based on their rating points. The batsmen are rated based on their performance in the match by adding new factors (pitch effect and home advantage) in the models. The proposed ranking is compared with the ICC's traditional ranking of batters in the test cricket series.

## 1. Introduction

Test cricket is a game of patience [1, 2], and it is littered with renowned batsmen who have set extremely high standards [3]. A batsman has nearly limitless time to set and play each ball individually. It is a ball-and-bat duel that is not defined by the number of deliveries [4]. Despite this, batsmen have struggled to stay at the crease, as conditions and lapses in focus result in the loss of their wicket, particularly in the modern game. In general, traditional measures such as batting average are used to evaluate a batsman in test cricket [5, 6]. Players were awarded points based on the number of runs they score during the game. This time-honored method has several drawbacks [5]. The context of the contest in which the runs are scored is not revealed by the runs scored in the match or the average runs in the series. For instance, scoring 100 runs in a low-scoring contest is not the same as scoring 100 runs in a high-scoring match. Scoring 100 runs

in the first innings is not the same as scoring 100 runs in the fourth innings. This is due to the match's circumstances. Due to deterioration in the pitch, the pitch of the first inning is radically different from the pitch of the fourth inning. Traditional measures such as batting average, on the other hand, overlook such instances. In the same way, playing in Melbourne is not the same as playing in Qaddafi Stadium. The impact of performing in one's own country or overseas, as well as other factors such as bowling first or second and losing or not losing the toss, are all said to have an impact on cricket results [7]. In any sport, there are various approaches for determining who the top player or team is [8, 9]. You can give them points based on which team performs best. There are both technical and nontechnical approaches for evaluating players or teams. Differentiating points are used to determine the winner in several sports [10]. It is possible to model point difference, but we will not. We are attempting to develop a new statistical measure that will allow us to assess

batsman performance in the context of the contest in which runs are scored. We are concentrating on the outcomes of Test Cricket matches. We explored the extent to which a variety of characteristics, such as playing at home or away, batting or fielding first, and pitch condition, influence match outcomes. Initially, we will forecast Test match outcomes using a multinomial logistic regression model. These forecasts will subsequently be used to evaluate the batter's performance.

A large number of studies have attempted to focus on Test Cricket players in various ways around distinct theoretical frameworks [5, 11–17]. Recent research studies on cricket have highlighted the need of examining and comprehending prematch indicators such as toss, ground effects, home ground, and rating of both participating teams, among others [18–22]. Kimber and Hansford studied cricket batting strategy at various levels [23]. They showed how scoring rate, opposition bowling strength, and pitch condition can be accurately integrated with runs scored to create an overall picture of batsmen's relative attributes. The Test match results were studied by Allsopp and Clarke [24]. They concluded that a team's first-inning bowling and batting strength, first-inning batting order lead, and home advantage are all good indicators of a winning test match outcome. Barooah and Mangan looked into some of the problems in evaluating batsmen for test matches [25]. They discovered that batters in cricket are mostly valued according to their average score: in test matches, an average of 50 or more provides a rule-of-thumb for distinguishing inordinate players from the purely good. Singh et al. assessed cricket players' batting performance and calculated the impact of their performance on the ICC ranking system [26]. Male test cricket batters and female test cricket batters were ranked by Rohde [12]. He proposed a straightforward approach for ranking batters based on their performance. Mukherjee used a diffusion-based PageRank algorithm on the networks to figure out how important it is to rate teams and captains [27]. In Test Cricket, Akhtar and Scarf predicted match results session by session [28]. They looked at how to match result probability (win, draw, and loss), and consequences differed from one session to the next. Daud and Muhammad collected a collection of Test matches [29]. They proposed a new ranking system for Test Cricket teams based on the number of runs scored and wickets taken. They suggested that a standard accuracy index be developed to determine the relevance of the discrepancy between the researcher's proposed rating system and the ICC rating system. Akhtar et al. developed a new rating system for players [5]. They determined the criteria for the best player in test cricket. Shah and Patel applied principal component analysis and weighted average method to rate the captain of captains among all 29 captains included in the study. Brewer and Stevenson suggested a survival analysis to forecast batting abilities in Test Cricket matches [30]. They developed a model in two stages, the first for individual players to assess their initial and balanced batting talents, as well as the rate of change in both. They matched and identified the cricketers who open the batting, which has a positive impact on the batting order. Hussain et al. utilized the International

Cricket Council's ad-hoc point system to assess cricket teams, and it is exclusively based on the number of wins and losses in cricket matches [31]. They compared their findings to those of the ICC. Boys and Philipson used an additive log-linear model to model run scores [13]. They looked at how an individual batsman's innings-by-innings variation in runs becomes a source of doubt in their ranking position. Stevenson and Brewer developed a Bayesian parametric model to calculate and estimate how intercontinental cricketers' batting ability alters across innings using a Gaussian process [32]. They identified which batsmen are struggling or improving their batting skills, which has a real-world influence on sportsman evaluation, aptitude recognition, and team selection strategy. Researchers have long hypothesized that the batter's performance influences the outcome of test matches [11]. In cricket, the concept of a player's rating appears to have always piqued the interest of sports analysts. The research on batsmen's performance also shows the importance of home ground, which can have a substantial impact on the outcome of a match.

## 2. Forecasting Test Matches' Results

All test cricket matches played between January 1, 2017, and December 31, 2019, will be considered. The cricket website ESPNcricinfo (<https://www.stats.cricinfo.com/ci/content/records/307847.html/>) is used to get session-by-session data. Rain-affected contests and those with poor lighting will be disqualified. A Test Cricket match lasts five days, with each day consisting of three sessions (lunch, tea, and end of the match). The study only included nine (out of ten) recent ICC (International Cricket Council) sanctioned Test Cricket playing countries. Afghanistan has been removed due to its current status as a Test-playing nation, and, as a result, its participation in a disproportionately small number of ICC-sanctioned matches. Outcomes are measured over three years since it is assumed that for the most part, the core playing group has stayed consistent throughout this time frame. At the end of each session, a series of multinomial nominal logistic regressions is fitted to forecast Test match outcome probabilities. Here, we will look at a model with a multinomial response (win, draw, and loss).  $Y$  depicts the match result by assigning values (1, 0, and  $-1$ ), with each value equating to a victory, a tie, or a defeat. The reference category is draw (0). We employed the Akaike information criteria (AIC) (Sakamoto, Ishiguro, and Kitigawa [33]), which is formulated as  $AIC = 2 * (\text{number of estimated parameters involved in the model}) - 2 * (\log\text{-likelihood})$  and Nagelkerke's  $R$  square to examine the model fit (Nagelkerke, [34]), which is given as  $R^2 = 1 - \exp\{-2(L_1 - L_0)/n\} / \{1 - \exp(2L_0)/n\}$ . In each session of each day, we modeled match outcome session-by-session in Table 1 and forecasted the test match outcome probabilities. In this section, we used those probabilities to assess each batter's contribution to both teams. To compare our suggested rating system to the existing batting average approach, three distinct Test match series (7 matches) were included. We display the rating points for covariates such as ground effect, no ground effect, home advantage, and no

TABLE 1: Session by session modeling.

Day	Session	Set of covariates	$K$	AIC	R_square (%)	Correct prediction (%)	Matches
	Start of the match	$g + h$	6	211.058	16.20	64.70	119
Day_1	Lunch	$g + h + l + w_1$	10	211.996	22.10	68.90	119
Day_1	Tea	$g + h + l + w_1$	10	202.106	32.10	68.10	119
Day_1	End	$g + h + l + w_1 + w_2$	12	192.97	39.20	73.10	119
Day_2	Lunch	$g + h + l + w_1 + w_2$	12	197.564	36.10	70.60	119
Day_2	Tea	$g + h + l + w_1 + w_2$	12	190.174	41.10	74.80	119
Day_2	End	$g + h + l + w_1 + w_2$	12	176.296	47.30	72.40	117
Day_3	Lunch	$g + h + l + w_1 + w_2$	12	162.248	57.00	81.20	117
Day_3	Tea	$g + h + l + w_1 + w_2$	12	155.36	56.70	81.40	113
Day_3	End	$g + h + l + w_1 + w_2$	12	142.556	57.30	77.70	103
Day_4	Lunch	$g + h + l + w_1 + w_2$	12	125.774	61.70	77.40	94
Day_4	Tea	$g + h + l + w_1 + w_2$	12	119.884	60.80	79.30	87
Day_4	End	$g + h + l + w_1 + w_2$	12	93.532	66.30	78.60	70
Day_5	Lunch	$g + h + l + w_1 + w_2$	12	78.166	64.60	80.00	45
Day_5	Tea	$g + h + l + w_1 + w_2$	12	47.012	85.10	83.30	30

home advantage. We use comparisons to see how these prematch factors affect batters' ratings.

### 3. Measuring Batters' Contribution

To determine the batter's contribution, you must first obtain the odds of the test match's outcome. Nominal multinomial logistic regression is used to calculate the match outcome probability (Sohail and Scarf, 2012). These actual probabilities are written as follows:

$$p^A(Y) = p_t(Y = y | H = h, G = g, L = l_t, W_1 = w_{1t}, W_2 = w_{2t}), \quad (1)$$

where  $P(Y)$  denotes the probability (win = 1, draw = 0, or loss = -1) at the end of each session  $t$  ( $t = 1st, 2nd, 3rd, \dots, 15th$ ),  $l$  denotes the lead until session  $t$ ,  $w_1$  denotes the first team's wickets,  $w_2$  denotes the second team's wickets,  $g$  denotes the ground effect, and  $h$  denotes home advantage. The model assumes  $Y$  has a multinomial distribution, that is,  $Y$  follows MN ( $p_{win}, p_{draw}, p_{loss}$ ) with,

$$\begin{aligned}
 p_{win} &= p\left(Y = \frac{1}{\text{covariates}}\right) = \frac{\exp(\alpha_{win} + \beta_{win}^T X)}{\{1 + \exp(\alpha_{win} + \beta_{win}^T X) + \exp(\alpha_{loss} + \beta_{loss}^T X)\}}, \\
 p_{draw} &= p\left(Y = \frac{0}{\text{covariates}}\right) = \frac{1}{\{1 + \exp(\alpha_{win} + \beta_{win}^T X) + \exp(\alpha_{loss} + \beta_{loss}^T X)\}}, \\
 p_{loss} &= p\left(Y = \frac{-1}{\text{covariates}}\right) = \frac{\exp(\alpha_{loss} + \beta_{loss}^T X)}{\{1 + \exp(\alpha_{win} + \beta_{win}^T X) + \exp(\alpha_{loss} + \beta_{loss}^T X)\}}.
 \end{aligned} \quad (2)$$

We forecast match outcomes based on the above-mentioned explanatory variables for each session of the test match. The potential position for both the reference team

and the opponent squad has also been well-defined. At the end of each session, the hypothetical position of the batsmen is defined as follows:

$$p^H(Y) = p_t(Y = y | H = h, G = g, L = l_t, W_1 = w_{1(t-1)}, W_2 = w_{2(t-1)}). \quad (3)$$

We assess their contributions after computing their points to determine the best batsman in the Test Cricket matches.

**3.1. Example 1.** Consider an Australia-New Zealand test match at the Perth Cricket Stadium in Australia. When we fit a model at the commencement of a test match, the likelihood

of the reference team (Australia) winning, drawing, and losing is 0.67, 0.03, and 0.30, respectively. Table 2 contains session-level data. Australia wins the match by 296 runs.

We depict the Trans-Tasman series, which is played in Australia between New Zealand and Australia. Australia has won this series (3-0). In the series, the Australian cricket team had the benefit of playing at home. The batters' rating points throughout the series are shown in Table 3. Table 4

TABLE 2: Scorecard from the 2019 Test Match between Australia (reference team) and New Zealand (opposite team) at Perth (ID 2374).

	Day_1			Day_2			Day_3			Day_4		
	Lunch	Tea	End	Lunch	Tea	End	Lunch	Tea	End	Lunch	Tea	End
Lead	76	160	248	337	416	307	250	325	417	436	369	296
W1	2	2	4	6	10	10	10	11	16	19	19	19
W2	0	0	0	0	0	5	10	10	10	12	15	20

TABLE 3: Proposed rating of batters in test series played between Australia and New Zealand in 2019–2020.

Batters (Australia)	Batters (New Zealand)	Matches	Rating	Ranking
SPD Smith		3	0.204	1
M Labuschagne		3	0.152	2
TM Head		3	0.08	3
	GD Philipson	1	0.053	4
	T A Blundell	2	0.045	5
DA Warner		3	0.038	6
MS Wade		3	0.038	7
MA Starc		3	0.036	8
	C de	3	0.035	9
	BJ Watling	3	0.035	10
	MJ Santner	2	0.023	1
	TA Astle	1	0.022	12
JR Pattinson		2	0.018	13
TD Paine		3	0.009	14
PJ Cummins		3	0.009	15
	HM Nicholls		0.007	16
JA Burns		2	0.006	17
	TG Southee	3	0.004	18
	LRPL Taylor	2	0.003	19
NM Lyon		3	0.002	20
	N Wagner	3	0.002	21
	MJ Henry	3	0.002	22
	T A Boult	1	0.001	23
	TWM Latham	1	0	24
	JA Raval	3	0	25
	KS Williamson	2	0	26
Hazlewood		2	0	27
	WER Somerville	1	0	28
	LH Ferguson	1	0	29
	Grandhomme	1		30

shows the results of our proposed methodology when both ground effect and home advantage are taken out of the equation. Table 5 shows the batters’ rankings, which are based on traditional batter averages. Labuschagne of Australia received the highest average of 91.50 points. Second, our criteria assign a score to batters based on the probability of each session’s test match result. Instead of the contribution shown in the typical position, a batter who performs well in a critical scenario receives additional rating points. As it stands, batters who perform well against highly rated teams earn more points than batters who do well against average teams. M Labuschagne received the best batsman of the series award in the Trans-Tasman series, as per traditional ratings; he scored the most runs with the highest average. The outcome would be different if the batsman of the series award was awarded using our proposed criteria. At last, we found correlation between our proposed rating

TABLE 4: Rating of batters in Test series played between Australia and New Zealand in 2019–2020 (the covariates home advantage and ground excluded from analysis).

Batters (Australia)	Batters (New Zealand)	Matches	Rating	Ranking
	T A Blundell	2	0.298	1
M Labuschagne		3	0.246	2
DA Warner		3	0.165	3
SPD Smith		3	0.101	4
TM Head		3	0.07	5
	GD Philipson	1	0.068	6
TD Paine		3	0.052	7
MS Wade		3	0.05	8
JA Burns		3	0.046	9
	LRPL Taylor	3	0.045	10
	TWM Latham	3	0.044	11
	J A Raval	2	0.031	12
	TA Astle	1	0.029	13
	KS Williamson	2	0.028	14
	C de Grandhomme	3	0.021	15
	BJ Watling	3	0.013	16
MA Starc		3	0.012	17
	MJ Santner	2	0.01	18
JR Pattinson		2	0.01	19
PJ Cummins		3	0.007	20
	HM Nicholls	2	0.007	21
	TG Southee	2	0.005	22
NM Lyon		3	0.003	23
	N Wagner	3	0.003	24
	MJ Henry	1	0.002	25
	T A Boult	1	0.002	26
	WER Somerville	1	0	27
	LH Ferguson	1	0	28

system and ICC rating system with  $r=0.636$  and  $p$ -value = 0.000.

3.2. *Example 2.* Consider the 2019 Test match between Pakistan (batting first) and Australia (batting second) at Brisbane, Australia. We used the coefficients of several covariates to fit a model using sessional data. Pakistan’s chances of winning, drawing, and losing at the start of the match are 0.53, 0.13, and 0.34, respectively. Table 6 contains session by session data. Australia wins this match.

Another example is a Test Cricket match series between Australia and Pakistan that took place in Australia in 2019. The series was won 2-0 by Australia. The Test Cricket series is depicted in Table 7. Table 7 shows the results of the analysis when all predictors are considered, whereas Table 8 shows the results when home advantage and ground effect are not considered. All batters who have at least one chance to bat

TABLE 5: ICC ranking of batters in a Test series played between Australia and New Zealand in 2019–2020.

Batters	Matches	Innings	Runs	Average	Ranking
M Labuschagne	3	6	549	91.5	1
DA Warner	3	6	297	49.5	2
TA Blundell	2	4	172	43	3
TM Head	3	5	213	42.6	4
TD Paine	3	4	153	38.25	5
SPD Smith	3	6	214	35.67	6
G D Philips	1	2	52	26	7
JA Burns	3	6	155	25.83	8
LRPL Taylor	3	6	152	25.33	9
C de Grandhomme	3	6	148	24.67	10
MS Wade	3	5	119	23.8	11
TWM Latham	3	6	126	21	12
T D Astle	1	2	42	21	13
MA Starc	3	4	76	19	14
BJ Watling	3	6	105	17.5	15
HM Nicholls	2	4	61	15.25	16
KS Williamson	2	4	57	14.25	17
JA Raval	2	4	45	11.25	18
PJ Cummins	3	4	41	10.25	19
JL Pattinson	2	2	16	8	20
MJ Santner	2	4	32	8	21
TG Southee	2	4	24	6	22
N Wagner	3	6	32	5.33	23
TA Boult	1	2	8	4	24
NM Lyon	3	4	15	3.75	25
W E R Somerville	1	2	7	3.5	26
MJ Henry	1	2	3	1.5	27
LH Ferguson	1	2	1	0.5	28
Hazlewood	1	1	0	0	29

TABLE 6: Scorecard from the 2019 Test match between Pakistan (reference team) and Australia (opponent team) in Brisbane, Australia.

	Day_1			Day_2			Day_3			Day_4		
	Lunch	Tea	End	Lunch	Tea	End	Lunch	Tea	End	Lunch	Tea	End
Lead	57	125	240	140	45	-72	-155	-292	-276	-196	-72	-5
W1	0	5	10	10	10	10	10	10	13	15	16	20
W2	0	0	0	0	0	1	3	5	10	10	10	10

for their team in a test series are rated in Table 7. In Tables 7, 8, and 9, DA Warner, an Australian batter, was ranked best among batters from both teams. In Tables 7 and 8, he had different scores. When the ground impact and home factor are removed from the model, he loses some ranking points in Table 7. According to Table 9, DA Warner remained the greatest batter with the highest batting average based on the ICC’s basic average criteria. In the series, DA Warner was named batsman of the series. It is concluded that there exists a correlation between our proposed criteria and ICC criteria with  $r = 0.835$  and  $p\text{-value} = 0.001$ .

3.3. *Example 3.* Take, for example, a test match played at Chattogram in 2018 between Bangladesh (reference team) and Sri Lanka. We used coefficients for different explanatory variables to fit the model on session-by-session data. For the reference team (Bangladesh), the chances of winning, drawing, and losing are 0.73, 0.12, and 0.15, respectively. The

match has been called a draw. Table 10 contains session-by-session lead data.

Consider another two Test Cricket match series in Bangladesh in 2018 between Sri Lanka and Bangladesh to further investigate the proposed criteria. The series was won by Sri Lanka with a score of 2-1. Players’ batting performance in the series is described in Table 11. When all predictors are included and techniques are used, Table 11 is produced. According to the results, Sri Lankan batter BKG Mendis received the most points (0.173) and was ranked first among all batters. When the covariate home advantage was not taken into account, BKG Mendis came in second with 0.201 points in Table 12.

When the covariates home factor and ground factor are removed from the collection of predictors, Table 12 is generated. In Tables 11 and 13, the same batter takes the first place. Table 13 is created to rate the players’ batting performances in the Test series using traditional averages. According to Table 13, Sri Lankan batsman BKG Mendis

TABLE 7: Proposed rating of batters in Test series played between Pakistan and Australia in 2019.

Batters	Batters	Matches	Rating	Ranking
(Australia)	(Pakistan)			
DA Warner		2	0.167	1
	Asad Shafique	2	0.083	2
M Labuschagne		2	0.078	3
JA Burns		2	0.067	4
Mohammad Rizwan		2	0.062	5
	Shan Masood	2	0.051	6
	Yasir Shah	2	0.05	7
	Azhar Ali	2	0.045	8
	Babar Azam	2	0.036	9
	Naseem Shah	1	0.016	10
	Imran Khan	1	0.015	11
MS Wade		2	0.009	12
NM Lyon		1	0.009	13
TM Head		1	0.007	14
	Haris Sohail	1	0.006	15
	Iftikhar Ahmed	2	0.006	16
PJ Cummins		1	0.005	17
	Muhammad Abbas	1	0.004	18
SPD Smith		2	0.004	19
TD Paine		1	0.004	20
Shaheen Shah Afridi		2	0.003	21
MA Starc		1	0.003	22
JR Hazlewood		1	0.003	23
	Muhammad Musa	1	0.002	24
	Imam-ul-Hauque	1	0.001	25

TABLE 8: Batters' rating in Test series played between Pakistan and Australia in 2019. The effect of ground and home advantage excluded from analysis.

Batters	Batters	Matches	Rating	Ranking
(Australia)	(Pakistan)			
DA Warner		2	0.241	1
M Labuschagne		2	0.13	2
MS Wade		2	0.078	3
	Asad Shafique	2	0.064	4
JA Burns		2	0.059	5
	Shan Masood	2	0.057	6
	Azhar Ali	2	0.056	7
Muhammad Rizwan		2	0.052	8
	Yasir Shah	2	0.039	9
	Babar Azam	2	0.033	10
	Naseem Shah	1	0.011	11
	Imran Khan	1	0.011	12
SPD Smith		2	0.011	13
NM Lyon		1	0.009	14
	Iftikhar Ahmed	2	0.007	15
TM Head		1	0.005	16
PJ Cummins		1	0.005	17
	Haris Sohail	1	0.005	18
	Muhammad Abbas	1	0.004	19
TD Paine		1	0.003	20
MA Starc		1	0.003	21
Shaheen Shah Afridi		2	0.003	22
	Muhammad Musa	1	0.002	23
	Imam-ul-Haq	1	0.001	24
JR Hazlewood		1	0	25

TABLE 9: ICC ranking of batters in a Test series played between Australia and Pakistan, 2019.

Batters	Matches	Innings	Runs	Average	Ranking
DA Warner	2	2	489	244.5	1
M Labuschagne	2	2	347	173.5	2
Babar Azam	2	4	210	52.5	3
JA Burns	2	2	101	50.5	4
MS Wade	2	2	98	49	5
Yasir Shah	2	4	194	48.5	6
Mohammad Rizwan	2	4	177	44.25	7
Shan Masood	2	4	156	39	8
Asad Shafique	2	4	142	35.5	9
TM Head	2	1	24	24	10
SPD Smith	2	2	40	20	11
Azhar Ali	2	4	62	15.5	12
Muhammad Abbas	1	2	30	15	13
TD Paine	2	1	13	13	14
NM Lyon	2	1	13	13	15
Iftikhar Ahmed	2	4	44	11	16
Muhammad Musa	1	2	16	8	17
PJ Cummins	2	1	7	7	18
Imran Khan	1	2	10	5	19
MA Starc	2	1	5	5	20
JR Hazlewood	2	1	5	5	21
Haris Sohail	1	2	9	4.5	22
Naseem Shah	1	2	7	3.5	23
Shaheen Shah Afridi	2	4	11	2.75	24
Imam ul Haque	1	2	2	1	25

TABLE 10: Scorecard from the 2018 Test match between Bangladesh (reference team) and Sri Lanka (opponent team) in Bangladesh (ID 2295).

	Day_1			Day_2			Day_3			Day_4			Day_5		
	Lunch	Tea	End	Lunch	Tea	End	Lunch	Tea	End	Lunch	Tea	End	Lunch	Tea	End
Lead	120	250	374	467	463	326	218	97	9	-99	-192	-119	-13	81	107
W1	2	2	4	7	10	10	10	10	10	10	10	13	13	15	15
W2	0	0	0	0	1	1	1	3	3	4	7	10	10	10	10

scored the most runs (271) and had the highest batting average in the series. Different criteria are used to grade batters in a conventional rating, and each table has a different top hitter, and our criteria rate batters by summing their hitting performance.

#### 4. Discussion

The results of the analysis revealed that each outcome had a varied impact at various stages of a Test Cricket match. Explanatory variables such as home factor, ground effect, and team strength have an effect on outcomes at the start of a Test match, but this effect fades as the match progresses. Lead has a minor impact at the start of a test match, but it grows in importance as the match develops. During the match, the number of wickets is also significant. A Test Cricket match is made up of five days, each of which has three sessions, for a total of fifteen sessions in a five-day contest. Predictors fluctuate their effect on match results over the course of the five-day match; therefore, we measured all of these sessions one-by-one to anticipate the outcomes at each phase, making it easier for forecasters to

forecast on a specific position. Through the statistical analysis, a rating system for Test Cricket matches is presented in this study. Multinomial logistic regression is used to calculate Test match outcome probability. To extend the scope of this study, a larger data set with additional explanatory variables can be used. There is fluctuation in our suggested rating system at the start and end of the Test match. A larger dataset can be used to tackle this problem. In our rating method, batters' contributions are judged by the difference between the hypothetical probability and the observed probabilities for the first inning and the difference between the supposed probability and the observed probabilities for the second inning. Researchers can utilize a variety of ways to overcome issues relating to the batter's contributions in a reduced-scoring game. The study methodologies used to rate batters are fairly practical because the proposed rating system is based on session probability, which assesses a batter's performance in relation to his contribution to the match outcome. We found correlation (0.883) with  $p$ -value=0.001 between proposed criteria and the traditional criteria introduced by ICC.

TABLE 11: Proposed rating of batters in Test series played between Sri Lanka and Bangladesh in 2018.

Batters	Batters	Matches	Rating	Ranking
(Sri Lanka)	(Bangladesh)			
BKG Mendis		2	0.173	1
DM de Silva		2	0.141	2
	Mominul Haque	2	0.100	3
	Tamim Iqbal	2	0.085	4
ARS Silva		2	0.080	5
	Imrul Kayes	2	0.064	6
	Mahidy Hassan Miraz	2	0.041	7
MD Gunathilaka		1	0.040	8
	Mahmudullah	2	0.040	9
	Mushfiqur Rahim	2	0.021	10
LD Chandimal		2	0.019	11
A Dananjaya		1	0.017	12
FDM Karunaratne		2	0.016	13
MDK Perera		2	0.016	14
	Liton Das	2	0.012	15
RAS Lakmal		2	0.009	16
	Sunzamul Islam	1	0.008	17
N Dickwella		2	0.005	18
	Mosaddek Hossain	1	0.003	19
	Abdul Razzaqk	1	0.002	20
	Tijul Islam	2	0.002	21
CBRLS Kumara		1	0.001	22
HMRKB Herath		2	0.001	23
	Sabbir Rahman	1	0.000	24
	Mustafizur Rahman	2	0.000	25
Sandakan		1	0.000	26

TABLE 12: Rating of batters in Test series played between Sri Lanka and Bangladesh in 2018 (the effect of home factor and ground excluded from analysis).

Batters	Batters	Matches	Rating	Ranking
(Bangladesh)	(Sri Lanka)			
Mominul Haque		2	0.222	1
	BKG Mendis	2	0.201	2
	DM de Silva	2	0.172	3
	ARS Silva	2	0.127	4
Tamim Iqbal		2	0.116	5
Imrul Kayes		2	0.092	6
Mehidy Hassan Mirza		2	0.068	7
Mushfiqur Rahim		2	0.067	8
	MD Gunathilaka	1	0.063	9
	LD Chandimal	2	0.053	10
Liton Das		2	0.045	11
Mahmudullah		2	0.033	12
	MDK Perera	2	0.033	13
	FDM Karunaratne	2	0.03	14
Sunzamul Islam		1	0.027	15
	N Dickwella	2	0.025	16
	RAS LAKmal	2	0.025	17
	A Dananjaya	1	0.023	18
	HMRKB Herath	2	0.008	19
Mosaddek Hossain		1	0.004	20
Abdul Razzaqk		1	0.004	21
	CBRLS Kumara	1	0.003	22
Tijul Islam		2	0.003	23
Mustafizur Rahman		2	0.001	24
Sabbir Rahman		1	0.000	25
Sandakan		1	0.000	26



TABLE 13: ICC ranking of batters in a Test series played between Sri Lanka and Bangladesh in 2018.

Batters	Matches	Innings	Runs	Average	Ranking
BKG Mendis	2	3	271	90.33	1
Mominul Haque	2	4	314	78.5	2
ARS Silva	2	3	235	78.33	3
DM de Silva	2	3	220	73.33	4
LD Chandimal	2	3	117	39	5
Mahmudullah	2	4	134	33.5	6
Liton Das	2	4	131	32.75	7
Mushfiqur Rahim	2	4	120	30	8
Tamim Iqbal	2	4	99	24.75	9
N Dickwella	2	3	73	24.33	10
Sunzamul Islam	1	1	24	24	11
Imrul Kayes	2	4	95	23.75	12
MDK Perera	2	3	70	23.33	13
Mehidy Hassan Mirza	2	3	65	21.67	14
MD Gunathilaka	1	2	30	15	15
FDM Karunaratne	2	3	35	11.67	16
RAS Lakmal	2	3	34	11.33	17
A Dananjaya	1	2	20	10	18
HMRKB Herath	2	3	26	8.67	19
Mosaddek Hossain	1	2	16	8	20
Mustafizur Rahman	2	3	13	4.33	21
Tijul Islam	2	3	8	2.67	22
CBRLS Kumara	1	1	2	2	23
Abdul Razzak	1	2	3	1.5	24
Sabbir Rahman	1	2	1	0.5	25
Sandakan	1				26

## Data Availability

Datasets are derived from public resources website (<http://www.espnccricinfo.com>) and made available with the article.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

The authors would like to acknowledge Prince Sultan University and the EIAS: Data Science and Blockchain Laboratory for their valuable support. Also, the authors would like to acknowledge the support of Prince Sultan University for Article Processing Charges (APC) of this publication.

## References

- [1] A. Edgar, *The Death of Test Cricket*, pp. 127-128, Taylor & Francis, Oxfordshire United Kingdom, 2020.
- [2] T. Chatterjee, S. Raizada, R. Chatterjee, S. Saha, B. Sarkar, and S. Bardhan, "Evolution of test cricket in the last two decades," *Annals of Tropical Medicine and Public Health*, vol. 23, no. 17, pp. 231-749, 2020.
- [3] G. Barr and R. Van den Honert, "Evaluating batsmen's scores in test cricket," *South African Statistical Journal*, vol. 32, no. 2, pp. 169-183, 1998.
- [4] T. Walle, *A Passion for Cricket: Masculinity, Ethnicity and Diasporic Spaces in Oslo*, Unpublished PhD Thesis: University of Oslo, Norway, 2010.
- [5] S. Akhtar, P. Scarf, and Z. Rasool, "Rating players in test match cricket," *Journal of the Operational Research Society*, vol. 66, no. 4, pp. 684-695, 2015.
- [6] O. G. Stevenson and B. J. Brewer, "Finding your feet: a Gaussian process model for estimating the abilities of batsmen in test cricket," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 70, no. 2, pp. 481-506, 2021.
- [7] A. M. Nevill and R. L. Holder, "Home advantage in sport," *Sports Medicine*, vol. 28, no. 4, pp. 221-236, 1999.
- [8] B. Gerrard, "A new approach to measuring player and team quality in professional team sports," *European Sport Management Quarterly*, vol. 1, no. 3, pp. 219-234, 2001.
- [9] M. Tavana, F. Azizi, F. Azizi, and M. Behzadian, "A fuzzy inference system with application to player selection and team formation in multi-player sports," *Sport Management Review*, vol. 16, no. 1, pp. 97-110, 2013.
- [10] D. L. Gill and T. E. Deeter, "Development of the sport orientation questionnaire," *Research Quarterly for Exercise & Sport*, vol. 59, no. 3, pp. 191-202, 1988.
- [11] I. P. Wickramasinghe, *Predicting the Performance of Batsmen in Test Cricket*, 2014.
- [12] N. Rohde, "An "economic" ranking of batters in test cricket," *Economic Papers: A journal of applied economics and policy*, vol. 30, no. 4, pp. 455-465, 2011.
- [13] R. J. Boys and P. M. Philipson, "On the ranking of test match batsmen," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 68, no. 1, pp. 161-179, 2019.
- [14] P. Scarf, X. Shi, and S. Akhtar, "On the distribution of runs scored and batting strategy in test cricket," *Journal of the Royal Statistical Society: Series A*, vol. 174, no. 2, pp. 471-497, 2011 Apr.
- [15] P. Scarf and S. Akhtar, "An analysis of strategy in the first three innings in test cricket: declaration and the follow-on," *Journal of the Operational Research Society*, vol. 62, no. 11, pp. 1931-1940, 2011 Nov.
- [16] P. Scarf, X. Shi, and S. Akhtar, "Modelling batting strategy in test cricket," *InProgress in Industrial Mathematics at ECMI 2008 2010*, pp. 481-489, Springer, Berlin, Heidelberg.
- [17] S. Akhtar, *Statistical Modelling in Test Cricket*, Doctoral dissertation, University of Salford.
- [18] K. P. Jayalath, "A machine learning approach to analyze ODI cricket predictors," *Journal of Sports Analytics*, vol. 4, no. 1, pp. 73-84, 2018.
- [19] A. Desai and B. Maharaj, "South Africa's Race to Return to Global Sport: Results and Prospects on home-ground—the Case of Cricket," *Sport in Society*, vol. 23, no. 1, 2019.
- [20] K. Kapadia, H. Abdel-Jaber, F. Thabtah, and W. Hadi, "Sport Analytics for Cricket Game Results Using Machine Learning: An Experimental Study," *Applied Computing and Informatics*, 2020.
- [21] A. Sinha, *Application of Machine Learning in Cricket and Predictive Analytics of IPL 2020*, Lambert academic publishing, Saarbrücken, Germany, 2020.
- [22] P. Tekade, K. Markad, A. Amage, and B. Natekar, "Cricket match outcome prediction using machine learning," *International Journal*, vol. 5, no. 7, 2020.
- [23] A. C. Kimber and A. R. Hansford, "A statistical analysis of batting in cricket," *Journal of the Royal Statistical Society: Series A*, vol. 156, no. 3, p. 443, 1993.

- [24] P. E. Allsopp and S. R. Clarke, "Rating teams and analysing outcomes in one-day and test cricket," *Journal of the Royal Statistical Society: Series A*, vol. 167, no. 4, pp. 657–667, 2004.
- [25] V. K. Borooah and J. E. Mangan, "The "Bradman class": an exploration of some issues in the evaluation of batsmen for test matches, 1877-2006," *Journal of Quantitative Analysis in Sports*, vol. 6, no. 3, 2010.
- [26] G. Singh, N. Bhatia, and S. Singh, "Fuzzy cognitive maps based cricket player performance evaluator," *International Journal of Enterprise Computing and Business Systems*, vol. 1, no. 2, pp. 1–15, 2011.
- [27] S. Mukherjee, "Identifying the greatest team and captain—a complex network approach to cricket matches," *Physica A: Statistical Mechanics and Its Applications*, vol. 391, no. 23, pp. 6066–6076, 2012.
- [28] S. Akhtar and P. Scarf, "Forecasting test cricket match outcomes in play," *International Journal of Forecasting*, vol. 28, no. 3, pp. 632–643, 2012.
- [29] A. Daud and F. Muhammad, "Ranking cricket teams through runs and wickets," in *Proceedings of the International Conference on Active Media Technology*, Maebashi, Japan, October 2013.
- [30] P. Shah and M. N. Patel, "Ranking the cricket captains using principal component analysis," *Int. J. Physiol. Nutr. Phys. Educ.*, vol. 3, no. 2, pp. 477–483, 2018.
- [31] A. Hussain, Y. Qiang, M. Abdul, K. Wu, Z. Zhao, and B. Ahmed, "Region-wise ranking for one-day international (ODI) cricket teams," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 10, 2019.
- [32] O. G. Stevenson and B. J. Brewer, "Bayesian survival analysis of batsmen in Test cricket," *Journal of Quantitative Analysis in Sports*, vol. 13, no. 1, pp. 25–36, 2017.
- [33] Y. Sakamoto, M. Ishiguro, and G. Kitagawa, "Akaike information criterion statistics," *D. Reidel*, vol. 81, no. 10, p. 26853, Dordrecht, The Netherlands, 1986.
- [34] N. J. D. Nagelkerke, "A note on a general definition of the coefficient of determination," *Biometrika*, vol. 78, no. 3, pp. 691–692, 1991.