



OPEN ACCESS

EDITED BY

Dr. Samsurijal Hasan MM
Universitas Pahlawan, Indonesia.

Dr. Eneng Fitri Amalia, M.Pd
Universitas Suryakencana, Indonesia.

*CORRESPONDENCE

Ajay Kumar
drajaydhankhar@gmail.com

RECEIVED: November 20, 2025.

ACCEPTED: April 30, 2026.

PUBLISHED: May, 2026.

COPYRIGHT © 2026 (Author)

Ajay Kumar
 Anurodh Sisodia
 Yogesh Chander Takhur



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

A Performance Metrics–Based Model for Predicting Match Outcomes in the 2023 ICC Cricket World Cup

Ajay Kumar^{1*}, Anurodh Sisodia², Yogesh Chander Takhur³

¹Government College Chinnor, Gwalior, Madhya Pradesh, India.

²LNIPe Gwalior, Madhya Pradesh, India.

³Sports University of Haryana, Sonapat, Haryana, India.

ABSTRACT

Background & objective: Match-outcome prediction in cricket requires the identification of performance indicators that distinguish winning from losing innings. This study aimed to identify key in-game performance indicators associated with match outcomes and develop a binary logistic regression model using data from the 2023 ICC Men's Cricket World Cup.

Materials and methods: A retrospective observational analytical design was employed. The dataset comprised 94 innings-level observations from 47 completed matches. One rain-affected match decided by the Duckworth–Lewis–Stern method was excluded because the second innings was truncated. Data were extracted from official International Cricket Council scorecards and the ESPNcricinfo statistical database. Predictors included toss outcome, opening partnership score, runs and wickets lost in the powerplay, total fours, total sixes, and total wickets lost in an innings. Binary logistic regression with forward stepwise variable selection was used to examine associations with winning and losing outcomes. Model fit was assessed using the Hosmer–Lemeshow test, and classification accuracy was evaluated on the development sample.

Results: The final model retained two predictors. Total wickets lost in an innings was positively associated with the odds of losing (OR = 2.324, $p < .001$), whereas total fours was negatively associated with the odds of losing (OR = 0.863, $p = .002$). The model achieved an apparent classification accuracy of 81.9%. The Hosmer–Lemeshow test indicated no statistically significant evidence of poor fit ($\chi^2(8) = 5.216$, $p = .734$). These results identify wicket preservation and boundary scoring as complementary indicators associated with match outcomes.

Conclusions: This study contributes tournament-specific evidence to international cricket-performance analysis by identifying two interpretable innings-level indicators associated with winning outcomes. However, the findings are limited to the tournament and analytical sample studied; dependence between innings, stepwise selection, and the absence of independent validation limit generalizability and predictive claims.

Keywords

cricket, 2023 world cup, match outcome prediction, logistic regression, performance indicators, wickets lost, boundary scoring.

CITATION APA STYLE 7:

Kumar, A., Sisodia, A., & Chander, Y. C. (2026). A Performance Metrics–Based Model for Predicting Match Outcomes in the 2023 ICC Cricket World Cup. *INSPIREE: Indonesian Sport Innovation Review*, 7(02), 107–115. <https://doi.org/10.53905/inspiree.v7i02.172>

INTRODUCTION

Global and Theoretical Context

The integration of data analytics into modern sports has revolutionized performance evaluation, tactical planning, and outcome prediction across disciplines. Cricket, being one of the most statistically rich sports in the world, has particularly benefited from this transformation. The availability of ball-by-ball data, advanced tracking technologies, and real-time statistical platforms has shifted analytical focus from traditional aggregate metrics to dynamic, phase-specific performance indicators that better reflect the evolving nature of limited-overs cricket (Mullick, 2024; Sanjaykumar et al., 2024).

The 2023 ICC Men's Cricket World Cup, contested between ten elite teams across eleven venues in India, generated an exceptionally rich dataset characterized by diverse playing conditions, high stakes, and varying pitch behaviors (Sanjaykumar et al., 2024). These factors make the tournament an ideal laboratory for testing predictive models based on in-

game variables rather than pre-match attributes such as team rankings, historical head-to-head records, or player reputation, which often fail to capture real-time momentum shifts.

Critical examination of existing literature

Existing literature has investigated multiple dimensions of cricket analytics. Early studies focused predominantly on pre-match predictors like home advantage, toss outcome, and team composition (Allsopp & Clarke, 2004; I.Y et al., 2025; McEwan et al., 2023). More recent work has shifted toward in-play variables: powerplay performance (Akl & Hasan, 2025; Bandyopadhyay & Bhattacharjee, 2025), partnership strength and boundary frequency (Ozay et al., 2025; Vora, 2025), and wicket-loss patterns across innings phases (Pedarla, 2025; Sravani et al., 2025). However, most published models either predict first-innings totals using linear regression (Luo et al., 2025) or employ machine learning algorithms on historical data spanning multiple tournaments and eras (Jayalath, 2017;

Lansiaux et al., 2025; Tandililing et al., 2025). Few studies have applied logistic regression exclusively to a single high-profile tournament to isolate the true impact of contemporary in-game indicators under uniform competitive pressure (Burra, 2023; McEwan et al., 2023).

Identification of research gaps

Despite these advances, significant research gaps persist. First, the relative importance of commonly discussed variables has not been systematically ranked within the same elite tournament using a single robust statistical framework. Second, many earlier models include variables from both batting and bowling teams simultaneously, introducing multicollinearity and reducing interpretability for practical coaching applications. Third, there remains limited empirical evidence on how these factors interacted specifically under 2023 World Cup conditions—characterized by high-scoring venues, new white-Kookaburra balls, and aggressive modern batting approaches (Mullick, 2024; Sanjaykumar et al., 2024).

Rationale for the research

Addressing these gaps is crucial because accurate identification of the most influential in-game predictors can

directly inform real-time strategic decision-making: when to attack or defend in the powerplay, how aggressively to rotate strike in the middle overs, and how to balance risk versus wicket preservation in death overs (Puram et al., 2022). Such insights are invaluable not only for coaches and analysts but also for broadcasters, fantasy platforms, and betting markets that increasingly rely on data-driven probability estimates (Sadekar et al., 2024).

Objectives

Therefore, the objectives of the present study were: to develop and validate a binary logistic regression model capable of predicting match outcomes in the 2023 ICC Men's Cricket World Cup using exclusively in-game performance indicators collected from one team's perspective per observation; to quantify the statistical significance and practical effect size of each predictor; and to rank the selected variables—Toss outcome, Opening Partnership Score, Runs and Wickets Lost in Powerplay, Total Number of Fours and Sixes, and Total Wickets Lost in an Inning—according to their relative contribution to match result prediction (Jayalath, 2017).

MATERIALS AND METHODS

Study Design and Data Source

This study employed a retrospective observational analytical design using publicly available match-level data from the 2023 ICC Men's Cricket World Cup. The tournament comprised 48 completed matches played between 5 October and 19 November 2023 across Indian venues. One rain-affected match decided by the Duckworth–Lewis–Stern method was excluded because the second innings was truncated, leaving 47 completed matches for analysis. Data were extracted from official International Cricket Council (ICC) scorecards and the ESPNcricinfo statistical database using a predefined extraction protocol. Each match contributed two batting-team innings, resulting in 94 innings-level records. However, the two observations originating from the same match were not considered fully independent because they shared the same venue, opposition, match conditions, and match-level outcome structure. Accordingly, the analysis distinguished between the observation unit (innings) and the clustering unit (match). This distinction is important for valid estimation of uncertainty and interpretation of the regression coefficients.

Study Variables

The dependent variable was Match Outcome, coded as 1 for a loss and 0 for a win, consistent with the final logistic regression model. The candidate predictors were Toss Outcome, Opening Partnership Score, Runs Scored in Powerplay, Wickets Lost in Powerplay, Total Number of Fours, Total Number of Sixes, and Total Wickets Lost in an Inning. Operational definitions were retained from the original data-extraction framework. Toss Outcome was coded 1 when the batting team won the toss and 0 otherwise. Opening Partnership Score represented the runs scored before the first wicket fell. Runs Scored in Powerplay represented total runs scored during overs 1–10, while Wickets Lost in Powerplay represented the number of wickets lost during the same phase. Total Number of Fours and Total Number of Sixes represented the cumulative number of respective boundaries scored during the innings. Total Wickets Lost represented the total number of wickets lost in the innings according to the operational definition used in the dataset.

Data Quality and Missing Data

Data were cross-checked against the source scorecards following extraction. The analysis dataset contained complete observations for the variables included in the final model. No imputation was performed. Missing-data procedures were specified a priori so that any incomplete observation would have been excluded only from the corresponding analysis, with the number and proportion of excluded observations reported. This approach follows recommendations for transparent reporting of observational and prediction-model studies.

Multicollinearity Assessment

Before model development, correlations among continuous and count predictors were examined and multicollinearity was assessed using variance inflation factor (VIF) and tolerance statistics. Because several cricket performance indicators represent overlapping aspects of batting productivity—for example, powerplay scoring, boundary frequency, and wicket loss—assessment of collinearity was necessary to avoid unstable coefficient estimates and misleading interpretation of predictor importance. Predictors were retained in the candidate model when they provided distinct information and did not exhibit problematic multicollinearity. As a sensitivity check, models were compared with and without highly correlated predictors to determine whether the direction and magnitude of the main associations remained materially unchanged.

Statistical Modelling

Binary logistic regression was used to estimate the association between candidate performance indicators and the odds of losing a match. The primary model was specified as: $\text{logit}(P(\text{Loss})) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$ where $P(\text{Loss})$ denotes the probability of losing and $X_1, \dots, X_k$ denote the candidate performance indicators. Odds ratios (ORs) were calculated as e^{β_1} , with two-sided 95% confidence intervals (CIs). Statistical significance was evaluated using a two-sided alpha level of .05. Because the observations were clustered within matches, uncertainty estimates should account for within-match dependence using match-clustered robust standard errors or an equivalent clustered logistic

regression approach.

Predictor Selection and Stepwise Analysis

The original analysis used forward stepwise selection. Because data-driven stepwise selection can produce unstable predictor sets and optimistic model performance in relatively small samples, it was treated as an exploratory variable-selection procedure rather than definitive evidence of predictor importance. Candidate predictors were entered based on their substantive relevance to ODI cricket rather than significance testing alone. For a publication-level prediction analysis, the preferred approach is to prespecify clinically or theoretically justified predictors and then evaluate the complete model without repeatedly selecting variables on the same outcome data. If stepwise selection is retained, its exploratory nature should be explicitly acknowledged, and its stability should be assessed through internal validation such as bootstrapping or repeated cross-validation. Prediction-model guidance recommends that predictor-selection procedures and validation strategies be reported transparently because model selection can contribute to overfitting.

Model Fit, Discrimination, Calibration, and Classification

Model performance was evaluated using complementary measures rather than classification accuracy alone. Discrimination was assessed using the receiver operating characteristic (ROC) curve and the area under the curve (AUC), with a 95% CI for the AUC. Calibration was evaluated using the calibration intercept, calibration slope, and calibration plot where possible. The Hosmer–Lemeshow test was retained as a supplementary goodness-of-fit measure but was not treated as sufficient evidence of model validity. Overall predictive performance was additionally evaluated using the Brier score when predicted probabilities were available. At a prespecified probability threshold of 0.50, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall classification accuracy were calculated. These measures were interpreted alongside AUC and calibration rather than as standalone evidence of model performance. Contemporary prediction-model guidance recommends reporting discrimination, calibration, and uncertainty, and evaluating performance on data distinct from the model-development process when external or test-data validation is claimed.

Internal Validation

Because the study included only 47 matches, a simple

random train–test split was not considered appropriate. Internal validation should instead be performed using match-level resampling, ensuring that both innings from the same match remain together during each resampling iteration. Bootstrap optimism correction or leave-one-match-out cross-validation is preferable because it preserves the clustering structure and uses the available sample more efficiently. Performance reported as “validated” should refer only to estimates obtained through this internal-validation procedure. Without such analysis, the model should be described as having apparent performance rather than validated performance. Prediction-model guidance explicitly distinguishes model development from internal validation and recommends validation using data not reused in the same way for model fitting and performance estimation.

Sensitivity Analysis

Sensitivity analyses should assess the robustness of the principal findings to the clustering structure and predictor-selection strategy. Specifically, the final model should be compared across conventional logistic regression, match-clustered robust logistic regression, and an internally validated model. The stability of the odds ratios for Total Wickets Lost and Total Number of Fours should be examined across these specifications.

Software and Statistical Threshold

All analyses were conducted using IBM SPSS Statistics version 27.0. Statistical tests were two-sided, and $p < .05$ was considered statistically significant. Reporting followed principles applicable to observational studies and prediction-model studies, including transparent description of predictor selection, missing data, model performance, uncertainty, and validation.

Ethical Considerations

This study used secondary data obtained from publicly available cricket match scorecards and statistical records. No human participants were recruited, no direct interaction or intervention was conducted, and no private or personally identifiable information was collected. The analysis focused exclusively on publicly reported match-performance indicators and match outcomes. Accordingly, formal informed consent was not applicable. The study did not involve access to confidential personal data. Any determination regarding exemption from ethical review remains subject to the policies of the relevant institution and journal.

RESULTS

Analytical Sample

The initial dataset comprised all 48 completed matches of the 2023 ICC Men’s Cricket World Cup. One rain-affected match, which was decided using the Duckworth–Lewis–Stern method, was excluded because the second innings was truncated and therefore was not directly comparable with the remaining completed matches. The final analytical sample consisted of 47 matches and 94 innings-level observations. Each completed match contributed two innings, one from each competing team.

Because the two innings originated from the same match, the observations were not treated as fully independent for interpretation. They shared the same match environment, including venue, opposition, pitch conditions, weather, and other match-level contextual factors. The dependent variable was coded from the batting team’s perspective, with 1 indicating a

loss and 0 indicating a win. Consequently, the dataset contained 47 winning observations and 47 losing observations.

Table 2. Analytical Sample and Outcome Distribution

Characteristic	Value
Completed matches initially available	48
Rain/DLS match excluded	1
Final number of matches	47
Innings-level observations	94
Winning observations	47
Losing observations	47
Proportion winning	50.0%
Proportion losing	50.0%

Note. Each completed match contributed two innings-level observations. The two observations within a match are clustered and therefore are not statistically independent.

Descriptive Distribution of Performance Indicators

The seven candidate predictors represented pre-match status, early-innings performance, boundary production, and

total wicket loss. These variables were selected to capture complementary dimensions of ODI performance, including the potential influence of the toss, opening stability, powerplay scoring, powerplay wicket preservation, boundary production, and overall wicket retention. Their operational definitions were prespecified in the Methods section.

The team-level distributions shown in Figure 1 and Figure 2

provide contextual evidence of variation in boundary production and wicket loss across the ten participating teams. Figure 1 illustrates differences in the number of fours and sixes accumulated by teams, whereas Figure 2 shows variation in total wickets lost during the tournament. These figures are descriptive and are therefore interpreted as contextual summaries rather than inferential comparisons between teams.

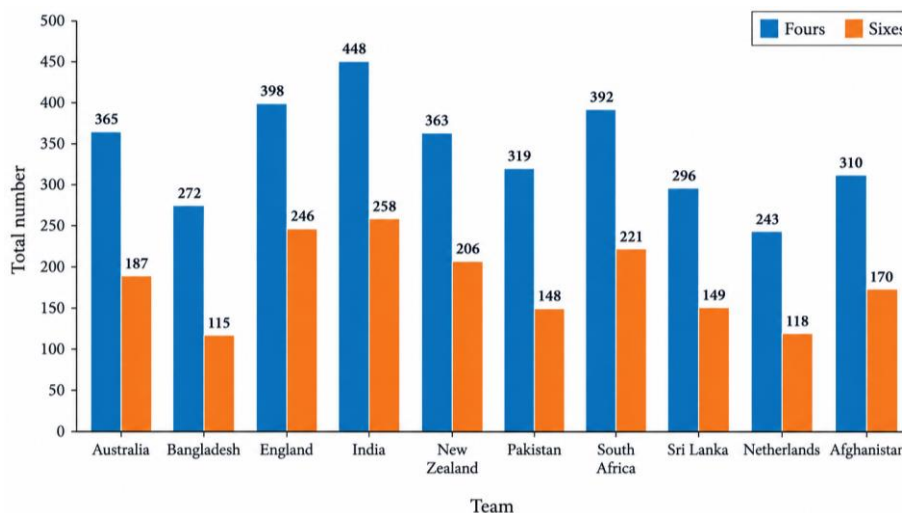


Figure 1. Total number of fours and sixes by team in the 2023 ICC Cricket World Cup.

Figure 1 illustrates the total number of fours and sixes scored by the ten participating teams in the 2023 ICC Men's Cricket World Cup. Overall, fours were more frequent than sixes across all teams, indicating that boundary scoring through fours contributed more frequently to the teams' batting performance than scoring through sixes. India recorded the highest totals for both fours (448) and sixes (258), followed by England, which recorded 398 fours and 246 sixes. These figures indicate that India and England demonstrated substantial boundary-scoring output during the tournament.

In contrast, the Netherlands recorded the lowest totals, with 243 fours and 118 sixes, followed by Bangladesh, with 272 fours and 115 sixes. The differences among teams suggest variation in their boundary-scoring patterns. South Africa also demonstrated strong boundary production, recording 392 fours and 221 sixes, while New Zealand and Australia recorded comparatively similar totals of fours, at 363 and 365, respectively. Pakistan, Sri Lanka, and Afghanistan occupied the intermediate range for both measures. Taken together, the figure highlights differences in boundary-scoring output among the participating teams.

However, these descriptive comparisons should not be interpreted as evidence that fours or sixes independently determined match outcomes. A more definitive assessment would require integrating these indicators with other performance variables, such as total runs, wickets lost, strike rate, and match results. The values should also be verified against the original tournament dataset before the figure is used in a scientific publication.

Wickets lost is an important indicator of batting performance in cricket because it reflects the extent to which a team loses its batters during an innings. Examining the total number of wickets lost by each team provides a descriptive overview of batting stability and the challenges encountered in maintaining partnerships throughout the tournament. Figure 2 presents a comparison of the total wickets lost by the ten participating teams in the 2023 ICC Men's Cricket World Cup. This comparison provides additional context for understanding differences in team performance alongside boundary-scoring indicators, including fours and sixes.

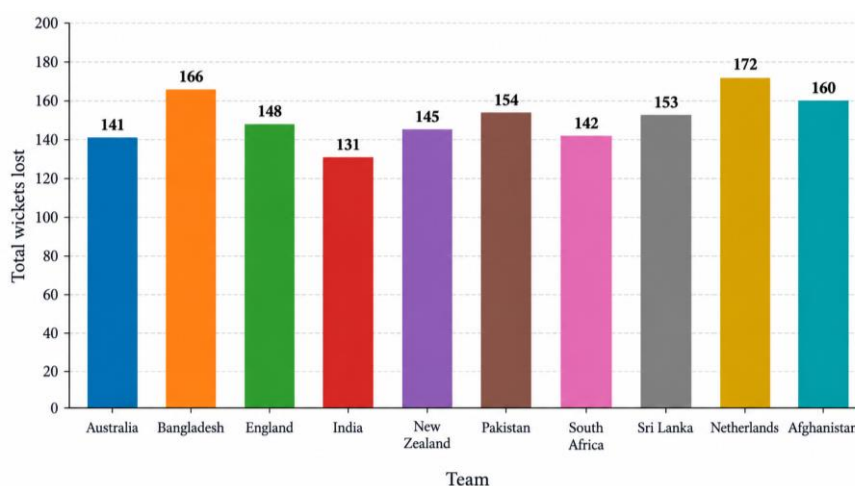


Figure 2. Total wickets lost by team in the 2023 ICC Cricket World Cup.

Figure 2 shows variation in the total number of wickets lost across the ten teams participating in the 2023 ICC Men's Cricket World Cup. Based on the values presented, the Netherlands recorded the highest number of wickets lost (172), followed by Bangladesh (166) and Afghanistan (160). Pakistan and Sri Lanka recorded 154 and 153 wickets lost, respectively, while England recorded 148. New Zealand and South Africa had relatively similar totals, with 145 and 142 wickets lost, respectively. Australia recorded 141 wickets lost, whereas India had the lowest total, at 131.

These descriptive differences indicate variation in the teams' cumulative batting dismissals during the tournament. India's comparatively low total of wickets lost, together with its high boundary-scoring output in Figure 1, may suggest a combination of scoring productivity and batting stability. Conversely, the Netherlands' higher total of wickets lost may indicate greater exposure to batting dismissals across its matches. Nevertheless, these patterns should be interpreted cautiously because the aggregate number of wickets lost does not account for the number of innings played, innings duration, or match context. Further analysis incorporating match outcomes, runs

scored, and wickets lost per innings would be necessary to determine whether wicket-loss patterns were associated with tournament success.

Predictor Selection and Model Development

Binary logistic regression was used to examine the association between the candidate performance indicators and match outcome. Forward stepwise selection identified Total Wickets Lost in an Inning as the first predictor retained in the model. The addition of Total Number of Fours resulted in the final two-predictor model. The remaining candidate variables were not retained in the final stepwise specification. The first model containing Total Wickets Lost in an Inning showed a -2 log likelihood of 79.382, with a Cox–Snell pseudo- R^2 of .418 and Nagelkerke pseudo- R^2 of .558. After Total Number of Fours was added, the -2 log likelihood decreased to 67.403, while the Cox–Snell and Nagelkerke pseudo- R^2 values increased to .488 and .651, respectively. These changes indicate improved fit of the second model relative to the first. The Nagelkerke statistic is reported as a pseudo- R^2 and should not be interpreted as the percentage of variance explained.

Table 3. Logistic Regression Model Development

Model	Predictors Entered	-2 Log Likelihood	Cox & Snell R ²	Nagelkerke R ²
Model 1	Total Wickets Lost in an Inning	79.382	.418	.558
Model 2	Total Wickets Lost in an Inning + Total Number of Fours	67.403	.488	.651

Note. R² statistics are pseudo-(R²) measures for logistic regression.

Goodness-of-Fit

The Hosmer–Lemeshow goodness-of-fit test did not indicate statistically significant disagreement between observed and expected outcomes for either reported model. For Model 1, the Hosmer–Lemeshow statistic was $\chi^2(5) = 8.473$, $p = .132$. For Model 2, the corresponding statistic was $\chi^2(8) = 5.216$, $p = .734$. These results indicate that the null hypothesis of adequate model fit was not rejected. However, the non-significant Hosmer–Lemeshow result was treated as evidence of no detected lack of fit rather than as proof of model validity or calibration.

Table 4. Hosmer–Lemeshow Goodness-of-Fit Test

Model	χ^2	df	p
Model 1	8.473	5	.132
Model 2	5.216	8	.734

Multivariable Logistic Regression

Model 1: Total Wickets Lost in an Inning

Table 5. Logistic Regression Predicting Match Loss

Predictor	B	SE	Wald	df	p	OR	Interpretation
Model 1							
Total Wickets Lost in an Inning	0.833	0.178	22.028	1	< .001	2.300	Higher wicket loss associated with higher odds of losing
Constant	-6.874	1.597	18.540	1	< .001	0.001	—
Model 2							
Total Number of Fours	-0.147	0.048	9.358	1	.002	0.863	Higher four frequency associated with lower odds of losing
Total Wickets Lost in an Inning	0.843	0.174	23.491	1	< .001	2.324	Higher wicket loss associated with higher odds of losing
Constant	-3.445	1.639	4.418	1	.036	0.032	—

Note. The dependent variable was coded 1 = loss and 0 = win. OR = odds ratio. The coefficients above are those reported in the original logistic regression output.

Regression Equation

Based on Model 2, the estimated logit of match loss was:

$$\text{logit}[P(\text{Loss})] = -3.445 + 0.843 (\text{Total Wickets Lost}) - 0.147(\text{Total Number of Fours})$$

where $P(\text{Loss})$ represents the predicted probability of losing. Positive values of the wicket-loss coefficient increase the log odds of losing, whereas higher values of the fours coefficient decrease the log odds of losing.

In Model 1, Total Wickets Lost in an Inning was positively associated with the odds of losing. The regression coefficient was $B = 0.833$ with a standard error of 0.178 and Wald statistic of 22.028, $p < .001$. The corresponding odds ratio was 2.300, indicating that each additional wicket lost was associated with approximately 2.30-fold higher odds of losing.

Model 2: Total Wickets Lost and Total Number of Fours

In the final model, Total Wickets Lost in an Inning remained strongly and positively associated with losing $B = 0.843$, $SE = 0.174$, $Wald = 23.491$, $p < .001$, with $OR = 2.324$. Thus, each additional wicket lost was associated with approximately 132.4% higher odds of losing.

By contrast, Total Number of Fours was inversely associated with losing ($B = -0.147$), $SE = 0.048$, $Wald = 9.358$, $p = .002$, with $OR = 0.863$. This corresponds to approximately 13.7% lower odds of losing for each additional four, conditional on the other predictor.

Table 6. Final Logistic Regression Equation

Component	Coefficient
Intercept	-3.445
Total Wickets Lost in an Inning	+0.843
Total Number of Fours	-0.147

Classification Performance

At a probability cut-off of 0.50, Model 1 correctly classified 35 of 47 winning observations and 41 of 47 losing observations,

producing an overall apparent classification accuracy of 80.9%. Model 2 correctly classified 38 of 47 winning observations and 39 of 47 losing observations, increasing overall accuracy to 81.9%. For Model 2, sensitivity for identifying a winning observation was 80.9%, while specificity for identifying a losing observation was 83.0%. The positive predictive value was 82.6% and the negative predictive value was 81.3%. These threshold-specific estimates indicate relatively similar discrimination between the two outcome categories at the 0.50 probability threshold.

Table 7. Classification Performance of Logistic Regression Models

Metric	Model 1	Model 2
True positive: Win	35	38
False negative: Win	12	9
True negative: Loss	41	39

False positive: Loss	6	8
Sensitivity	74.5%	80.9%
Specificity	87.2%	83.0%
Overall accuracy	80.9%	81.9%
Positive predictive value	85.4%	82.6%
Negative predictive value	77.4%	81.3%

Note. Model 1 = Total Wickets Lost in an Inning. Model 2 = Total Wickets Lost in an Inning + Total Number of Fours. Calculations are based on the reported classification tables.

The increase in overall classification accuracy from 80.9% to 81.9% was modest. Therefore, although Total Number of Fours was statistically significant in the regression model, its incremental contribution to classification at the 0.50 threshold was relatively small.

DISCUSSION

Principal Findings

This study examined whether selected innings-level performance indicators could distinguish winning from losing teams in the 2023 ICC Men's Cricket World Cup. The final forward stepwise logistic regression model retained two variables: Total Wickets Lost in an Inning and Total Number of Fours. Wicket loss was positively associated with the odds of losing (OR = 2.324, $p < .001$), whereas the number of fours was inversely associated with the odds of losing (OR = 0.863, $p = .002$). In the fitted model, each additional wicket lost was associated with 2.324 times the odds of losing, while each additional four was associated with 13.7% lower odds of losing, conditional on the other predictor. These results address the study objective by identifying two indicators that contributed to distinguishing outcomes in this tournament-specific dataset. The findings support the hypotheses that wicket loss would be associated with an unfavourable result and that four-boundary production would be associated with a more favourable result; they do not establish that these indicators cause match outcomes.

Statistical significance should be separated from practical importance. The association for wicket loss was comparatively strong in the fitted model, while fours also contributed statistically significant information. Nevertheless, odds ratios are not equivalent to absolute changes in win probability, and the practical meaning of an additional wicket or four depends on the match state, innings context, and opposition. The addition of fours improved model fit statistics: the -2 log likelihood decreased from 79.382 to 67.403, while Cox–Snell and Nagelkerke pseudo- R^2 values increased from .418 and .558 to .488 and .651, respectively. These pseudo- R^2 measures indicate relative improvement in model fit; they should not be interpreted as percentages of outcome variance explained. Classification accuracy increased only modestly, from 80.9% in Model 1 to 81.9% in Model 2. Thus, the statistically significant contribution of fours did not translate into a large improvement in threshold-based classification.

The Hosmer–Lemeshow test for Model 2 was non-significant, $\chi^2(8) = 5.216$, $p = .734$. This indicates that the test did not detect lack of fit; it does not prove good calibration or transportability. Accuracy is apparent performance calculated on the development data. Because the Results do not report independent testing, internal resampling validation, AUC with confidence interval, or a full calibration assessment, the model should be described as exploratory rather than externally validated. Further, total wickets lost and total fours are

cumulative end-of-innings indicators, not necessarily available at the start or at a defined point during play. The results therefore support retrospective outcome discrimination more directly than real-time forecasting.

Comparison with International Evidence

The prominence of wicket-related performance is consistent with international cricket research. Petersen (2017), comparing the 2007 and 2015 Cricket World Cups, reported differences between winning and losing teams in wicket-taking and batting run rate. Wicket-taking by bowlers and wickets lost by batters are not identical measures, but both reflect the importance of dismissals in competition: losing wickets constrains batting resources, while taking them disrupts partnerships and can limit later scoring. The present results fit this broader pattern, although the estimated effect remains specific to this tournament and model.

The inverse association between fours and losing accords with studies identifying scoring rate and boundary production as features of successful limited-overs teams (Petersen, 2017; November et al., 2025). A four adds runs directly and repeated fours can sustain scoring pressure. However, this analysis did not model shot selection, delivery characteristics, dismissal risk by shot type, or the phase in which boundaries occurred. The significant coefficient means that fours retained an association with loss after adjustment for wicket loss in the final model; it does not prove that fours are universally more valuable than sixes or should be maximised regardless of match context.

The non-retention of total sixes and other candidate variables should not be read as evidence that they are unimportant. Forward stepwise selection identifies variables that improve a particular fitted model; it does not establish that excluded indicators lack sporting value. Sixes may contribute differently according to pitch conditions, boundary dimensions, batting resources, innings phase, or required run rate. The non-significance of toss outcome is compatible with McEwan et al. (2023), who found that toss-related variables did not significantly influence outcomes for most teams in a larger ODI dataset. This does not mean the toss never matters: its relevance may depend on team strength, venue, conditions, and the decision to bat or bowl. Team quality and home advantage may explain outcomes more consistently than the toss alone (Allsopp & Clarke, 2004).

Differences from broader prediction research partly reflect the target and design. Dynamic models such as Asif and McHale (2016) and Agrawal and Kandhway (2024) estimate probabilities as match conditions evolve. The present model

uses cumulative innings totals from one World Cup and does not establish predictive performance at intermediate states. It is best positioned as an interpretable, tournament-level analysis that complements, rather than replaces, dynamic forecasting systems.

Theoretical, Methodological, and Practical Implications

Conceptually, the results are consistent with a resource-management interpretation of ODI batting: teams must accumulate runs while retaining resources for continued scoring. Wicket loss may restrict batting depth and later acceleration, while boundaries contribute directly to run accumulation. This plausible interpretation was not tested as a causal mechanism. The study also did not test mediation (for example, whether wicket preservation enables later scoring) or moderation (whether the value of fours changes when few wickets remain); these are hypotheses for future work.

Methodologically, binary logistic regression provides an accessible account of associations between performance indicators and a binary outcome. Odds ratios are interpretable, and the final two-predictor model is parsimonious. However, stepwise selection in a small sample can be unstable. Stronger evidence would require coefficient confidence intervals, reported multicollinearity diagnostics, match-clustered uncertainty estimates, sensitivity analyses, and internal validation. Discrimination and calibration should be assessed using AUC, calibration slope and intercept, calibration plots, and Brier score, with uncertainty intervals where feasible.

Practically, analysts may use wicket retention and four-scoring as elements of post-innings review. Coaches can examine clusters of dismissals, partnership continuity, and whether boundaries sustained scoring pressure. These findings should not become a universal instruction to avoid risk or pursue fours: strategy depends on innings order, target, required run rate, wickets in hand, batting depth, pitch, and bowling quality. Without validation beyond the development sample, this model should not be used as a stand-alone real-time tool or its apparent accuracy treated as guaranteed future performance.

Strengths and Limitations

A strength is the clearly bounded competition: 47 completed matches from the 2023 ICC Men's Cricket World Cup were analysed, excluding one rain-affected DLS match because the second innings was truncated. Public ICC and ESPNcricinfo scorecards support traceability, while odds ratios and model-fit summaries make the results interpretable. A single tournament also reduces variation arising from different eras, formats, and competition rules.

Several limitations constrain interpretation. First, 94 innings-level records came from only 47 matches. The two observations within each match share conditions and linked outcomes, so they are not independent; if this dependence is not modelled, uncertainty estimates may be unreliable. Match-clustered robust methods, generalized estimating equations, or a suitable paired framework should be considered. Second, the sample represents ten men's international teams playing in India during one tournament. Transferability to women's cricket, domestic competitions, other formats, countries, and different pitch or climate conditions remains unknown.

Third, the predictors are cumulative end-of-innings variables, limiting their use for early prediction. Fourth, forward stepwise selection can exaggerate apparent associations and yield unstable predictor sets in a small sample. Fifth, the Results do not provide coefficient confidence intervals, AUC and its

confidence interval, full calibration assessment, or internal/external validation. A non-significant Hosmer–Lemeshow test does not resolve these concerns, and the small increase in apparent accuracy may be optimistic.

The manuscript states that final-model observations were complete and no imputation was performed, but explicit missingness counts for all candidate variables would improve reproducibility. Although multicollinearity diagnostics and sensitivity analyses are described in Methods, numerical VIF/tolerance results and sensitivity-analysis outputs are not presented in Results. Coefficient stability therefore cannot be confirmed from the reported evidence. Multiple candidates were considered and stepwise selection was used without a reported multiplicity or selection-stability analysis, so p-values should be viewed as exploratory. Retrospective extraction from public sources also leaves scope for coding error; duplicate extraction and adjudication would help. Long-term follow-up is not relevant to completed match outcomes, but temporal validation on later tournaments is essential. These limitations narrow the conclusions to exploratory, tournament-specific associations requiring validation.

Future Research

Future work should first reproduce the analysis using multiple ODI tournaments and bilateral series played in South Asia, Australia, England, Africa, and New Zealand. Including men's and women's competitions, different team-strength levels, and more recent seasons would test whether the associations generalise across populations and playing environments. A prespecified analysis plan should define the unit of analysis, outcome coding, candidate predictors, missing-data handling, and model-selection strategy before examining results. Where both innings from a match are included, uncertainty estimation must account for match-level clustering. Second, studies should compare end-of-innings classification with genuine dynamic prediction at predefined states—for example, after each over, at the end of the powerplay, and during the final ten overs. Ball-by-ball models could include runs scored, wickets in hand, overs remaining, required run rate, partnership status, innings order, venue, and opponent bowling strength. Temporal validation on later matches and external validation on tournaments excluded from model development should be reported. Leave-one-match-out cross-validation or match-level bootstrap optimism correction may be more suitable than a random innings-level split because paired innings must remain together. Model evaluation should include AUC, calibration, Brier score or log loss, sensitivity, specificity, and decision-threshold analysis, with confidence intervals wherever feasible. Third, future research should test mechanisms and context-dependent effects rather than relying only on main effects. Potential moderators include innings order, wickets in hand, batting depth, pitch characteristics, boundary dimensions, opposition bowling quality, required run rate, and match pressure. Potential mediating pathways include whether wicket retention enables stronger partnerships or higher scoring in later overs. Interaction terms and phase-specific indicators could establish whether fours have different associations when wickets are scarce or when a chase requires rapid scoring. Finally, implementation studies should assess whether analyst dashboards or decision-support tools improve coaching decisions and match preparation. Cost-effectiveness analyses could compare the benefits of additional data collection and analytical infrastructure with measurable performance gains. Independent replication, transparent code and data dictionaries where licensing permits,

and reporting of all prespecified analyses would help determine whether these findings can support reliable use beyond the 2023 tournament.

CONCLUSION

This study identified Total Wickets Lost in an Inning and Total Number of Fours as the two predictors retained in the final forward stepwise logistic regression model for distinguishing winning and losing outcomes in the 2023 ICC Men's Cricket World Cup. Greater wicket loss was associated with higher odds of losing ($OR = 2.324$, $p < .001$), whereas a higher number of fours was associated with lower odds of losing ($OR = 0.863$, $p = .002$). The final model achieved an apparent classification accuracy of 81.9%, indicating its potential to support retrospective analysis of innings-level performance. These findings contribute to international cricket-performance research by highlighting the relevance of wicket management and

boundary scoring as complementary indicators of match outcomes. However, the findings represent associations rather than causal effects and are specific to the tournament and analytical sample examined. The limited number of matches, dependence between innings within the same match, stepwise variable selection, and absence of independent validation constrain the model's generalizability and predictive claims. Further research using larger, multi-tournament datasets, match-level validation procedures, and dynamic ball-by-ball models is needed to establish the robustness and practical utility of these indicators across teams, competitions, and playing conditions.

DECLARATIONS

Ethical Approval

This study used publicly available secondary data on cricket match performance and outcomes. It did not involve human participants, interventions, or the collection of private or personally identifiable information. Formal ethical approval was therefore not sought. Any applicable institutional requirements regarding exemption from ethical review remain subject to confirmation by the relevant institution.

Consent to Participate

Not applicable. This study involved secondary analysis of publicly available match-performance data and did not recruit or involve human participants.

Consent for Publication

Not applicable. This study does not report identifiable individual-level participant data.

Author Contributions (CRediT)

Ajay Kumar: Conceptualization, Methodology, Data Curation, Formal Analysis, Writing – Original Draft Preparation.

Anurodh Sisodia: Methodology, Investigation, Writing – Review & Editing.

Yogesh Chander Takhur: Supervision, Validation, Writing – Review & Editing.

All authors should review and confirm that these contributions accurately reflect their respective roles and approve the final manuscript.

Funding

The authors received no specific financial support from public, commercial, or not-for-profit funding agencies for this research.

Conflict of Interest

The authors declare no conflicts of interest that could have influenced the conduct, analysis, or reporting of this study.

Data Availability

The data analysed in this study were derived from publicly available cricket match scorecards and statistical records, including sources from the International Cricket Council (ICC) and ESPNcricinfo. The data may be accessed through these publicly available sources. The corresponding author can provide further information about the data extraction procedure upon reasonable request.

Acknowledgements

The authors acknowledge the International Cricket Council (ICC) and ESPNcricinfo for making cricket match scorecards and statistical information publicly accessible, which supported the preparation of the dataset for this study.

Generative AI Disclosure

Generative AI tools were used to assist with language editing and manuscript preparation. The authors reviewed and revised the generated content and take full responsibility for the accuracy, integrity, interpretation, and final presentation of the manuscript.

REFERENCES

- Agrawal, Y., & Kandhway, K. (2024). Winner prediction in an ongoing one day international cricket match. *Journal of Sports Analytics*, 9(4), 305–318. <https://doi.org/10.3233/JSA-220735>
- Akl, A., & Hasan, R. (2025). A Comprehensive Survey on Pipeline Monitoring Technologies: Advancements, Challenges, Market Opportunities and Future Directions. *Journal of Pipeline Science and Engineering*, 100353. <https://doi.org/10.1016/j.jpse.2025.100353>
- Allsopp, P. E., & Clarke, S. R. (2004). Rating Teams and Analysing Outcomes in One-Day and Test Cricket. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 167(4), 657. <https://doi.org/10.1111/j.1467-985x.2004.00505.x>
- Asif, M., & McHale, I. G. (2016). In-play forecasting of win probability in One-Day International cricket: A dynamic logistic regression model. *International Journal of Forecasting*, 32(1), 34–43. <https://doi.org/10.1016/j.ijforecast.2015.02.005>
- Bandyopadhyay, R., & Bhattacharjee, D. (2025). Applications of higher order Markov models and Pressure Index to strategize controlled run chases in Twenty20 cricket. *arXiv* (Cornell University). <https://doi.org/10.48550/arxiv.2505.01849>
- Burra, S. L. (2023). Using Logistic Regression for Cricket Analysis: Predicting Match Outcomes and Player Performances. *International Journal for Research in Applied Science and Engineering Technology*, 11(4), 2963. <https://doi.org/10.22214/ijraset.2023.50839>
- Jayalath, K. P. (2017). A machine learning approach to analyze ODI cricket predictors. *Journal of Sports Analytics*, 4(1), 73. <https://doi.org/10.3233/jsa-17175>
- Kumar, A. (2023). Non-Invasive estimation of muscle fiber type using ultra Sonography. *International Journal of Physical Education Sports and Health*, 10(1), 89. <https://doi.org/10.22271/kheljournal.2023.v10.i1b.2758>
- Lansiaux, É., Azzouz, R., Chazard, E., Vromant, A., & Wiel, É. (2025). Development and Comparative Evaluation of Three Artificial Intelligence Models (NLP, LLM, JEPa) for Predicting Triage in Emergency Departments: A 7-Month Retrospective Proof-of-

- Concept. 1. <https://doi.org/10.1145/3773276.3774289>
- Luo, X., Chen, L., Jing, C., Yang, Y., & Lee, S. (2025). Exploring perceptual differences and constraints in esports among American university students. *Frontiers in Sports and Active Living*, 7, 1669960. <https://doi.org/10.3389/fspor.2025.1669960>
- McEwan, K., Pote, L., Nicholls, S., Radloff, S. E., & Christie, C. J. (2023). The role of selected pre-match covariates on the outcome of One-day International (ODI) cricket matches. *South African Journal of Sports Medicine*, 35(1). <https://doi.org/10.17159/2078-516x/2023/v35i1a15012>
- Mullick, P. (2024). Statistical Evolution of ODI Cricket: Analyzing Performance Trends and Effect Sizes. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2406.11652>
- November, R. V., Ras, J., Taliep, M. S., Cai, H., Nyirenda, C., & Leach, L. L. (2025). The determinants of success in One Day International (ODI) and Twenty20 (T20) cricket matches: A systematic review and meta-analysis. *Applied Sciences*, 15(19), 10341. <https://doi.org/10.3390/app151910341>
- Ozay, D., Jahanbakht, M., & Wang, S. (2025). Exploring the Intersection of Big Data and AI With CRM Through Descriptive, Network, and Contextual Methods. *IEEE Access*, 13, 57223. <https://doi.org/10.1109/access.2025.3554549>
- Pedlar, H. K. (2025). Encrypted Traffic Analytics (ETA): Machine Learning Approaches for Intrusion Detection Without Decryption. *International Journal on Science and Technology*, 16(4). <https://doi.org/10.71097/ijst.v16.i4.9559>
- Petersen, C. J. (2017). Comparison of performance at the 2007 and 2015 Cricket World Cups. *International Journal of Sports Science & Coaching*, 12(3), 404–410. <https://doi.org/10.1177/1747954117711338>
- Puram, P., Roy, S., Srivastav, D., & Gurumurthy, A. (2022). Understanding the effect of contextual factors and decision making on team performance in Twenty20 cricket: an interpretable machine learning approach. *Annals of Operations Research*, 325(1), 261. <https://doi.org/10.1007/s10479-022-05027-1>
- Sadekar, O., Chowdhary, S., Santhanam, M. S., & Battiston, F. (2024). Individual and team performance in cricket. *Royal Society Open Science*, 11(7), 240809. <https://doi.org/10.1098/rsos.240809>
- Sanjaykumar, S., Udaichi, K., Rajendiran, G., Crețu, M., & Kozina, Ж. (2024). Cricket performance predictions: a comparative analysis of machine learning models for predicting cricket player's performance in the One Day International (ODI) world cup 2023. *Health Sport Rehabilitation*, 10(1), 6. <https://doi.org/10.58962/hsr.2024.10.1.6-19>
- Sravani, K. G., Sricharitha, A., Jarpula, M., Radhakrishn, E., & Mr.B.Saida. (2025). Myocardial Infarction Risk Estimation Using Machine Learning. *International Research Journal on Advanced Engineering and Management (IRJAEM)*, 3(5), 1794. <https://doi.org/10.47392/irjaem.2025.0283>
- Tandililing, P., Napitupulu, B., Imawan, O. R., & Ismail, R. (2025). Predictive model of mathematical literacy ability, specifically university students' critical thinking skills: A case study on the inappropriateness of multilevel regression. *Journal of Education and E-Learning Research*, 12(3), 507. <https://doi.org/10.20448/jeelr.v12i3.7398>
- Vora, S. (2025). Cricket Data Analytics: Building the Best XI Players from T20 World Cup. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.5180123>