Beyond the Boundary: Revolutionizing the IPL MVP Index

Shreyas Rajaram - The Dalton School '26
Sehaj Doshi - Awty International School '25
Aarush Bahel - Haberdashers' Boys' School '25
Benjamin Pearlberg - The Pennington School '25
Neil Jaiswal - Bridgewater-Raritan Regional High School '25

Abstract

Since its inception in 2008, the Indian Premier League (IPL) has attracted many of the world's most skilled cricket players, offering a highly competitive arena for them to showcase their talents. Each season, the IPL awards the Most Valuable Player (MVP) title to the player who achieves the highest rating on the league's MVP metric. Ideally, this award recognizes the top performer of the season, with high rankings indicating outstanding achievement among elite players. However, the calculation used by the IPL to assess player performance lack consistency, are limited in scope, and rely on arbitrary criteria. This paper employs a multivariate regression model to develop a more robust formula, assigning mathematically optimized weights to devised metrics that better capture player contributions. With an R^2 value of 0.80—compared to the existing system's 0.66—this new formula provides a more accurate and comprehensive evaluation of player performance.

Introduction

The Indian Premier League (IPL) is a franchise-based T20 cricket league and currently ranks as the second most lucrative sports league globally, generating $9.5 billion in revenue in 2023.

Since its inception in 2008, the IPL has attracted many of the world's strongest cricketers, providing a highly competitive platform to showcase their abilities. In this talent-packed tournament, the Most Valuable Player (MVP) award holds immense significance as it is widely regarded as a prestigious honor among cricketers. Currently, the method of determining this award is arbitrary, limited in scope, and flawed in reasoning. This paper seeks to address these shortcomings by developing a new metric to improve the existing rating system.

Overview of Cricket Rules

The fundamental objective in T20 cricket, the format used in the IPL, is to score more runs than the opposing team within 20 overs (an over is a set of six balls delivered by a bowler/pitcher). There are three main ways to score runs: fours, sixes, and running between the wickets. Four runs are awarded when a batsman hits the ball past the boundary after it bounces at least once in play (similar to a ground-rule double in baseball). Six runs are awarded when the ball crosses the boundary on the fly (similar to a home run). Lastly, a run between the wickets occurs when a batsman hits the ball into play, and both batsmen swap ends; each swap counts as one run. Batsmen can swap ends multiple times, though typically do so once or twice to minimize the risk of being run-out if the ball reaches the stumps before they do. A batsman is out if the ball is caught on the fly by a fielder, they are run-out, or the ball hits the stumps behind them. When a batsman is out, they are replaced and cannot bat again in that game. If 10 out of 11 batsmen are out before all overs are bowled, the inning concludes. Refer to the appendix for additional details on cricket-specific terms and rules mentioned in this paper.

The Traditional MVP Metric: A Flawed System

Historically, the IPL's MVP award has been determined through metrics focused on individual match events, assigning points as follows:

- **Batting:** 3.5 points per six, 2.5 points per four
- **Bowling:** 3.5 points per wicket, 1 point per dot ball
- **Fielding:** 2.5 points per run-out or stumping, 2.5 points per catch

While these metrics offer a straightforward means of quantifying contributions, they fail to capture the full scope of player performance. For example, a batsman's ability to score runs quickly (strike rate) or a bowler's efficiency in conceding runs (economy rate) are not adequately emphasized, as accumulating boundaries and dot balls do not necessarily reflect these qualities. Furthermore, situations such as the following arise: for two batsmen who each face six balls, a batsman scoring 8 runs through 1s or 2s is credited with fewer MVP points than a counterpart who scores one single boundary and five dot balls, despite the former contributing more runs. Similarly, a bowler who concedes 20 runs in their allotted 4 overs with 4 dot balls would earn fewer MVP points than one who concedes 40 runs with 3 dot balls—an illogical outcome. These flaws could result in players who make significant contributions to their teams' success—without relying on boundaries, wickets, or dot balls—being undervalued.

Established Cricket Statistics

While cricket analytics are still in a relatively early stage of development, several statistics have traditionally been used to rate player performances. Although these metrics are not currently included in the IPL's MVP rating system, this paper will incorporate the following:

Batting Statistics

Runs Scored (RS): The total number of runs scored by a batsman in the season.

Strike Rate (SR): The speed at which a batsman scores runs.

$$SR = \frac{Runs\ Scored_{batter} \cdot 100}{Balls\ Faced_{batter}}$$

Batting Average (BA): The number of runs a batsman scores per dismissal.

$$BA = \frac{Runs\ Scored_{batter}}{Number\ of\ dismissals_{batter}}$$

Bowling Statistics

Balls Bowled (BB): The total number of balls bowled by a bowler in a season.

Runs Allowed (RA): Total runs conceded by a bowler in a season.

Economy (Econ): Measures the runs a bowler concedes per over.

$$Econ = \frac{Runs\ Conceded_{bowler}}{Overs\ Bowled_{bowler}}$$

Wickets (W): Wickets attributed to a bowler for a season.

Balls per Wicket (BpW): The average number of balls bowled per wicket taken.

$$\mathrm{B}p\mathrm{W} = \frac{BB_{bowler}}{W_{bowler}}$$

Methodology

From a general perspective, our model incorporates several devised metrics inspired by Strike Rate, Batting Average, Economy, and Balls per Wicket. These metrics are benchmarked against league averages for the season and adjusted based on the frequency of actions. Ultimately, individual player contributions are aggregated by team and correlated with the team's winning percentage for the season.

Statistics Developed for Our Model

In cricket, player evaluation metrics such as Strike Rate, Batting Average, and Economy provide insight into performance but often lack contextual weight. The devised metrics in this model address this gap by assessing batting and bowling efficiency in relation to team success, aiming to isolate individual contributions from overall team dynamics. Each formula has been constructed to capture distinct aspects of a player's efforts relative to the league standards. This approach enhances the model's ability to quantify a player's impact, especially under high variability conditions, such as match-to-match fluctuations in strike rates, dismissals, and bowler economies.

1. Hitting Value (HV):

$$HV = \frac{\left(SR_{batter} - SR_{avg}\right) \cdot RS_{batter}}{100}$$

The Hitting Value (HV) metric quantifies a batsman's ability to score efficiently, comparing their strike rate to the league average. A higher strike rate indicates faster scoring, which is critical for maximizing a team's total runs within a limited number of overs. The difference between an individual batsman's strike rate and the league average measures their scoring efficiency relative to others. Dividing by 100 sets this difference in runs per ball, as strike rate is calculated per 100 balls. Multiplying by the batsman's total runs scored scales the HV to reward players who sustain high efficiency over larger volumes of runs, thereby capturing both the speed and volume of scoring. This approach also penalizes batsmen with low strike rates, especially if they face a high number of deliveries without significant run production, as this hampers team momentum.

2. Dismissal Weighted Runs Above Average (dwRAA):

$$dwRAA = \left(BA_{batter} - BA_{avg}\right) \cdot dismissals_{batter}$$

dwRAA accounts for a batsman's consistency and output by factoring in their batting average relative to the league, scaled by the number of times they are dismissed. The metric compares the batsman's average against the league average, capturing their consistency in converting appearances into runs. Multiplying by the number of dismissals controls for batsmen who score heavily without frequent dismissals, ensuring that an exceptionally high average due to few dismissals does not unduly skew a player's evaluation. This normalization maintains balance by rewarding consistent scoring and penalizing those with inflated averages due to minimal dismissals.

3. Runs Below Average (RBA):

$$RBA = \left(\frac{ECON_{avg}}{6} \cdot BB_{bowler}\right) - RA_{bowler}$$

RBA assesses a bowler's ability to restrict scoring, a critical skill for maintaining team control over run rates in limited-overs cricket. The formula first divides the league's average economy rate by 6, converting the rate into runs conceded per over. By multiplying by balls bowled, an expected runs value can be derived based on league norms. Subtracting the actual runs allowed compares performances to league average: a positive RBA indicates that the bowler concedes fewer runs than expected, reflecting an ability to limit scoring. This metric effectively rewards bowlers who maintain control over opposing batsmen and sustain a low economy rate, both of which are essential in matches.

4. Wicket Frequency (WF):

$$WF = W_{bowler} \cdot \left(BpW_{avg} - BpW_{bowler}\right)$$

WF captures a bowler's effectiveness in taking wickets, which directly contributes to disrupting the opposition's batting lineup. The formula computes the frequency of wicket-taking by comparing the average balls per wicket across the league to the bowler's balls per wicket. Multiplying this ratio by the bowler's total wickets rewards bowlers who maintain efficiency while taking a high number of wickets. A high WF score is indicative of a bowler who consistently requires fewer deliveries to take wickets, constantly disrupting the opposing batting lineup and therefore limiting the opposition's scoring potential.

Each metric aligns with core principles of cricket strategy, where efficiency and consistency are valued over raw output. For batting, HV and dwRAA prioritize fast and consistent run-scoring, which are essential for accumulating a competitive total. Similarly, RBA and WF for bowling emphasize economy and wicket-taking frequency, both critical to curbing an opposing team's progress. By anchoring each metric to league averages, the model ensures fair comparisons among players with varying roles. It balances traditional measures of performance (strike rate, average, economy) with a more nuanced approach that reflects actual game impact.

The model also incorporates fielding, recognizing that taking catches and effecting run-outs and stumpings (stumpings are a specific type of run-out) are critical to a team's success. Given the lack of fielder tracking data to evaluate the difficulty of individual performances, the model uses two fielding statistics: the number of catches taken and the number of run-outs and stumpings effected. However, unlike the existing system's arbitrary weightings, the multivariate regression model accounts for the lesser impact of fielding performances on team success compared to batting and bowling metrics.

Data

The data for this paper was obtained from two sources: the Cricsheet Database and the official IPL website. The Cricsheet Database was used for play-by-play analysis, while the IPL website provided data on team performance per season, existing MVP ratings, and fielding statistics, including catches and run-outs. This model utilized data from the 2021 to 2024 IPL seasons, as fielding data has been consistently available only since 2021. Data from the 2022 to 2024 seasons was used for training, while the 2021 season was used for testing.

Multivariate Regression Model

For the training dataset, each of the six aforementioned metrics was Z-scored for every player in individual seasons and then aggregated for each team. The aggregated Z-scores were subsequently correlated with each team's winning percentage. The logic is that teams with stronger metric aggregates should win more matches overall, regardless of whether the contributions come from a large group of above-average players or a single exceptional performance. The optimized equation is as follows:

$$\widehat{\mathbf{Win}\,\%} = \left(\boldsymbol{a} \cdot \sum_{\boldsymbol{p=1}}^{\boldsymbol{Players}} HV_{z_p}\right) + \left(\boldsymbol{b} \cdot \sum_{\boldsymbol{p=1}}^{\boldsymbol{Players}} dwRAA_{z_p}\right) + \left(\boldsymbol{c} \cdot \sum_{\boldsymbol{p=1}}^{\boldsymbol{Players}} RBA_{z_p}\right) + \left(\boldsymbol{d} \cdot \sum_{\boldsymbol{p=1}}^{\boldsymbol{Players}} WF_{z_p}\right) + \left(\boldsymbol{f} \cdot \sum_{\boldsymbol{p=1}}^{\boldsymbol{Players}} Catches_{z_p}\right) + \left(\boldsymbol{g} \cdot \sum_{\boldsymbol{p=1}}^{\boldsymbol{Players}} (Stumpings + RunOuts)_{z_p}\right)$$

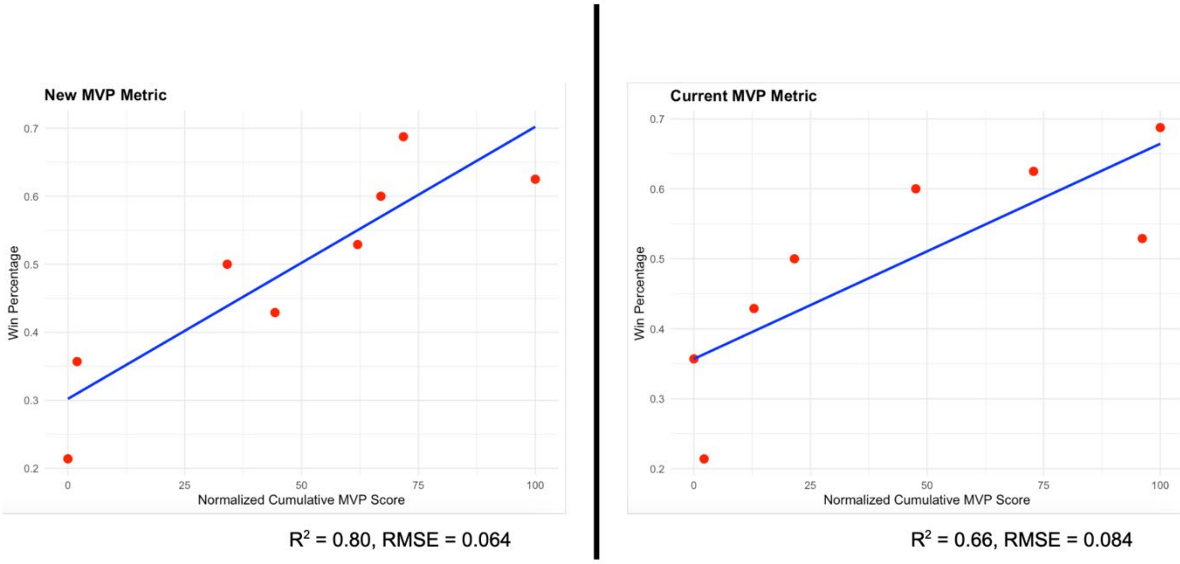
$$\boldsymbol{a = .007385} \qquad \boldsymbol{b} = .014514$$

$$\boldsymbol{c} = .017521 \qquad \boldsymbol{d} = .004933$$

$$\boldsymbol{f} = .002754 \qquad \boldsymbol{g} = .001485$$

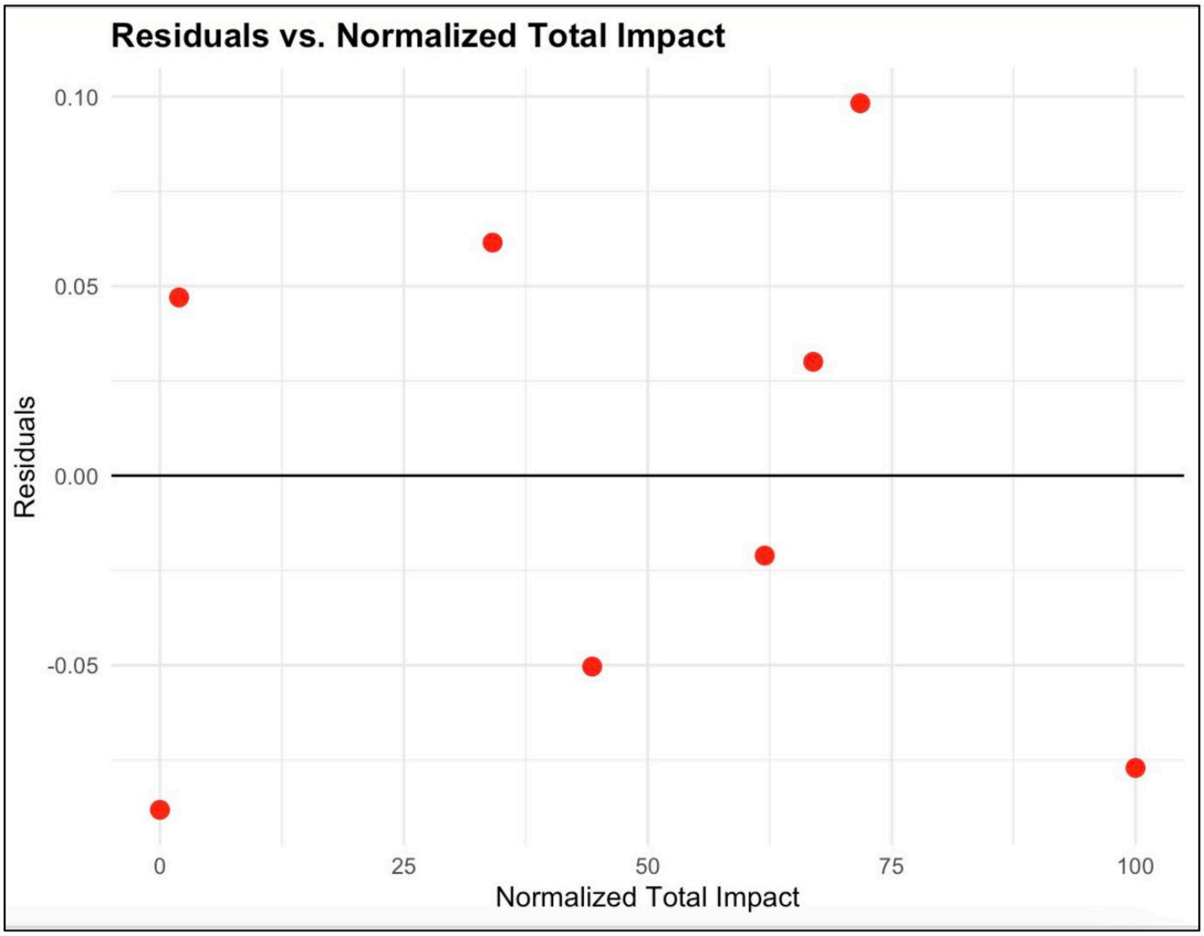
Ultimately, when applied to our testing data, this model had an R^2 value of 0.80, greater than the existing method's value of 0.66.

Correlation Graphs of IPL Data (2021-2024)



Overall, our correlation is significantly stronger than that of the current metric, even when evaluated using data from a single season.

Residual Plot

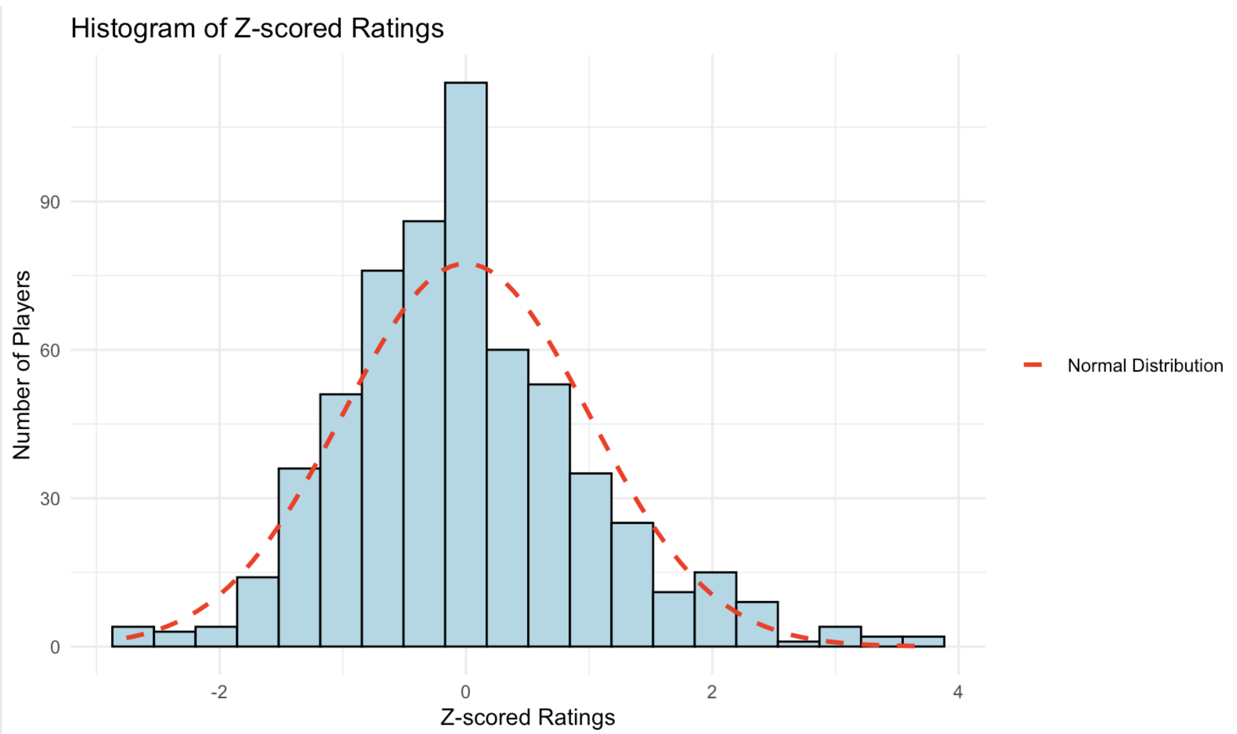


The residual plot above shows no visible trend, indicating that the regression model has no consistent systematic error.

Player MVP Calculation

Using the equation above, individual player ratings are calculated with a minor adjustment. To ensure that specialist batsmen or bowlers are not unfairly penalized for poor performance or limited opportunities in the opposite discipline, players with lower ratings in either batting or bowling (after the coefficients for dwRAA and HV or RBA and WF are applied and the metrics summed) are capped at a minimum rating of 0. In other words, a batsman with a positive batting rating (after considering Z-scored HV and dwRAA) and a negative bowling rating will have the bowling rating set to 0. However, for players with below-average performances in both disciplines, only the lower of the two ratings will be set to 0, while the other rating remains negative.

Player Rating Distribution



The graph above illustrates that the data is mostly normally distributed, suggesting that our rating system effectively captures a balanced distribution of both stronger and weaker players. Most players are clustered around the league average rating, corresponding to a Z-score of 0.

Discussion

The multivariate regression model proves to be a stronger predictor of team performance compared to the existing metric. Additionally, the optimized coefficients reveal several key implications, including:

- **Appropriate Weighting of Disciplines:** Both batting and bowling are given roughly equal weight, as the sums of the coefficients for the batting and bowling statistics used in the regression model are nearly equal. This balance is a strength of our model, as it reflects the necessity for a cricket team to perform well in both disciplines to win matches. Fielding, on the other hand, is given significantly less weight, as fielding contributions are one of many ways to dismiss batsmen. The model correctly places less importance on fielding compared to batting and bowling, reflecting the relative impact of these disciplines on team success.
- **Batting – Balance Between Volume and Efficiency:** For batting, the volume of runs a batsman scores is his most critical quality. However, the efficiency with which a batsman scores runs remains significant. According to our model, a batsman cannot be effective by excelling in only one of these aspects – he must perform well in both. This is logical, since a highly rated batsman should be able to score runs consistently without negatively impacting the team's momentum through slow run-scoring.
- **Bowling – Control Over Wickets:** There has been an age old debate in cricket that questions whether taking wickets or conceding fewer runs is more important. Our model suggests that bowlers who control the run rate are far more valuable than those who frequently take wickets. The coefficient for RBA is over 3.5 times greater than that of

WF, a disparity much larger than the difference between the batting statistics, where dwRAA is less than twice as significant as HV. Particularly in the IPL, where matches are played over 20 overs and teams are rarely all out, it is sensible that controlling runs is a greater contributor to winning matches than taking wickets.

Results

Comparison of the Top 10 Players Across Both Metrics

Top 10 Using Current MVP Metric

NAME	RATING	YEAR
Sunil Narine	100	2024
Jos Buttler	86.0	2022
Shubman Gill	76.1	2023
Rashid Khan	72.1	2023
Yashasvi Jaiswal	71.1	2023
Virat Kohli	70.0	2024
Mohammed Shami	67.9	2023
Cameron Green	66.2	2023
Hardik Pandya	63.1	2022
Faf du Plessis	62.8	2023

Top 10 Using Our Impact Rating

NAME	RATING	YEAR
Sunil Narine	100	2024
Shubman Gill	99.4	2023
Jos Buttler	91.2	2022
Jasprit Bumrah	89.4	2024
KL Rahul	85.5	2021
Faf du Plessis	83.6	2023
Faf du Plessis	81.5	2021
Virat Kohli	81.4	2024
Rashid Khan	77.5	2022
Ravindra Jadeja	76.8	2021

Ratings are scaled to 100

In the tables above, the top 10 players from the past four years have been normalized to a 100-point scale for both the existing MVP metric and our Impact Rating. The most notable difference between the results of our Impact Rating and the traditional MVP metric lies in the spread of the data. Using the traditional MVP metric, the rating gap between the third-place player and the best player is larger than the gap between the best and the tenth-best player in our Impact Rating. Since this list consists of established, world-class international cricketers, the compression in our results underscores a key strength of our method: its ability to provide a more balanced assessment across the player pool.

Our Impact Rating also highlights strong performances across a season that may have lacked the glamor of wickets and boundaries. For instance, Jasprit Bumrah, ranked 4th on our list, had an exceptional season as a bowler in 2024, which statistically was the most batting-dominated IPL season in history. However, his strength manifested less in taking wickets and more in restricting run-scoring, even as his counterparts struggled to do the same. While Bumrah does not even figure in the top 10 players under the traditional metric (having less than 60% of the impact of Sunil Narine), our model appropriately credits his performance with a rating of 90% of Narine's impact.

There are, however, some agreements between the two rating systems since strong performances often include scoring boundaries and taking wickets. The top three players are identical, featuring Sunil Narine, who excelled in 2024 as both a top batsman and bowler, as well as Jos Buttler and Shubman Gill, both of whom were the highest run-scores of their respective seasons by a considerable margin.

Case Studies

To demonstrate the efficacy of our new calculation compared to the existing MVP metric, this section examines the 2021 IPL season performances of two key players: Harshal Patel and AB de Villiers.

Harshal Patel:

Patel was named the MVP of the 2021 season for his record-breaking 32 wickets. However, his high economy rate, particularly in one of the lowest-scoring seasons ever, undermined his overall contribution. Despite his wicket-taking prowess, he conceded a significant number of runs per ball, reducing the overall effectiveness of his performance. Our model places a greater value on bowling control (RBA) than on wicket-taking (WF). While Patel's WF was among the highest of the season, it was offset by a negative RBA, ultimately lowering his overall rating. In our system, he ranked 44th for the season. This, while above average, reflects the detrimental impact of his high economy rate compared to his wicket-taking achievements.

AB de Villiers:

De Villiers was ranked 40th in the traditional MVP rankings in 2021. While his batting strike rate and average were above the league average, his lower total runs limited his MVP score. In contrast, our model accounted for the impact of his consistently high strike rate even with a relatively modest number of boundaries, as well as his fielding contributions — he was the top-rated fielder of the season. As a result, our model ranked him 8th for the 2021 season, reflecting a more comprehensive assessment of his all-round contribution.

Conclusion

The new MVP metric provides a more accurate representation of a player's contribution to their team's success by considering a broad range of factors and leveraging advanced statistical techniques to quantify individual performances. Compared to the existing rating system, it significantly improves the ability to capture a player's all-round impact.

Despite its strengths, the current model has room for improvement. While the context-independent nature of the metric makes it accessible and easy to interpret for players and fans, an ideal rating system would account for the context of player contributions. For example, our method treats a batsman scoring 40 runs off 20 balls in a losing effort the same as one achieving the same score while facing the opposition's best bowlers in a match-winning situation. Future iterations of the metric could integrate contextual factors, such as the strength of opposing players and the significance of key match moments, to deliver a more nuanced evaluation.

A significant limitation of cricket analytics lies in the availability of comprehensive data. The data used in this model was limited to play-by-play scoring data, which restricts the depth of analysis, particularly in areas such as strategy execution. Additionally, there is a scarcity of scholarly research on player impact valuation, with limited exploration into alternative performance metrics. As cricket analytics continues to evolve, more effective statistics and innovative data collection methods will emerge, offering exciting new possibilities for refining player evaluations.

Acknowledgements

We would like to thank Professor Adi Wyner from the Wharton Moneyball Academy for inspiring us to build this model. Additionally, we are grateful to Zeke Kelz, Jonathan Pipping, Ryan Brill, and Rafiz Sadique for their invaluable guidance throughout the development and presentation of our model.

Appendix: Glossary of Cricket Terms and Rules

Cricket to Baseball Dictionary:

- Bowler = Pitcher
- Batsman = Hitter
- 4s ~ ground rule doubles
- 6s ~ home runs
- Batting Average ~ Player runs scored per appearance
- Wickets/Dismissals = Outs
- Economy (ECON) ~ ERA per over (six balls)
- Run-out ~ Groundout
- Catch ~ Flyout
- Dot Balls ~ Strike

Batting:

Ways to score runs:

- Sixes (6s) - A batsman hits the ball over the boundary on the fly. (Comparable to a home run in baseball)
- Fours (4s) - A batsman hits the ball past the boundary after it bounces in play at least once. (Comparable to a ground-rule-double in baseball)
- Running Between the Wickets/Stumps - A batsman hits the ball into play, and both batsmen swap ends. Each successful swap counts as one run.

Batting Average (BA):

BA = (Number of runs a batsman scores) / (Number of times a batsman gets out)

Strike Rate (SR):

SR = (Number of runs a batsman scores) / (Number of balls a batsman faces) * 100

Bowling/Fielding:

Wickets (W): Number of batsmen a bowler dismisses (gets out)

Economy Rate (Econ):

Econ = (Number of runs a bowler concedes) / (Number of overs bowled)

Run-Outs: A run-out occurs when batsmen are running between the wickets and the fielding team gets the ball to one of the ends before a batsman has crossed the crease line near the wicket.

Dot Balls (Dot): A delivery (pitch) that does not result in any runs being scored. (Comparable to a strike in baseball)

References

Board of Control for Cricket in India. "MVP Rating." IPLT20. Last modified May 2024. Accessed November 8, 2024. https://www.iplt20.com/stats/2024.

Cricsheet. "Match Data." Cricsheet. Last modified November 8, 2024. Accessed November 8, 2024. https://cricsheet.org/matches/.

Pearlberg, Benjamin. "Could Moneyball Methodology Improve Indian Cricket?" Firstpost. Last modified August 2024. Accessed November 8, 2024. https://www.firstpost.com/opinion/could-moneyball-methodology-improve-indian-cricket-13806487.html.