

Involvement of Computing to improve decision making in Cricket

Jahangir Kamal, Manoj Kumar, Kailash Kumar, Aditya Tandon, Phong Thanh Nguyen

II. LITERATURE SURVEY

Abstract: In shorter format of Cricket, the choice of a bowler has three main parameters namely: economy, strike rate and dot balls delivered. In most of the cases, the most hitting parameters are economy rate and number of wickets taken, which again are inter related with the dot balls delivered. This paper presents a survey operational linear approach which comparative analyze the above-cited three parameters and suggests a solution based approach to choose a best bowler in "Playing Eleven" with highest preference to the dot balls delivered. The inter-relationship among these parameters are established based on collected data. The proposed indicator is proved useful while making decisions. A software-based architecture is also proposed relating to decision support system for selecting a bowler in playing eleven using past data.

Index terms: Twenty-twenty match, cricket, bowler selection, indicator, parameter, decision tree, bowling score.

I. INTRODUCTION

Importance of statistical modelling and data analysis in sports has been steadily increasing over the years. As teams have turned towards numbers to add substance alongside "field tactics", success stories in sports like Baseball, football and others have aplenty. Thankfully, Cricket is not too far from their counterparts either, with intense groundwork being done to improve. Main motivation for writing this paper is to make better decisions regarding selection of bowlers using statistics to recognize their strengths. Furthermore, this paper also aims to introduce other parameters, which can result in more discipline bowling.

Revised Manuscript Received on December 22, 2018.

Jahangir Kamal, Manoj Kumar, PhD Scholar, Jagannath University, Jaipur [kamaljhangirmit@gmail.com](mailto:kamaljahangirmit@gmail.com)

Kailash Kumar, Assistant Professor, CSE Department, Manav Rachna University, Faridabad

Aditya Tandon, Assistant Professor, College of Computing and Informatics, Saudi Electronic University, Riyadh, Kingdom of Saudi Arabia

Phong Thanh Nguyen, Research Scholar, Amity University, Noida, UP Director, Department of Project Management, Ho Chi Minh City Open University, Vietnam

As evolution of sports science continues, smaller details are observed and recorded. ^[1]Doljin and Fuss proposed for a smarter Cricket ball to record and calculate dynamic performances data, which can be stored in laptop or computer. ^[2]Laura Justham, Andrew West suggested

about the bowling system, which will be essential in recreating a normal delivery of ball, thus helping batsmen to train better for certain situation of a match.

Bowling action remains a field of interest, which has generated frequent studies. ^[3]Renshaw and Fairweather analyzed how batmen need to anticipate the upcoming delivery despite bowler's attempts to disguise it with his bowling actions. On medical grounds, ^[4]Burnett, Elliott and Marshall observed heart rate and blood lactate over duration of a 12 over spell. ^[5]Ranson, Brennett, King, Patel, O'Sullivan analyzed the risk of lower back injuries happening because of bowling actions. Given how physically demanding bowling for a considerable period can be, it is necessary to keep analyzing the constant effect on body as the game progresses. Not just restricted to bowling actions, ^[6]Chin identified some kinematic differences to identify the spin of the ball. However, related to scope of this paper, ^[7]Lemmer pointed out that batting average is important, however, other parameters like Strike Rate and consistency are equally important factors to be considered when it comes to judging quality of a batsman.

III. DATA ANALYSIS IN TWENTY OVER FORMAT

In contrast to other formats of Cricket, Twenty-twenty offers more thrill to the viewers. Owing to the limited number of overs there are to play for, each team tries to score as high as they can. This results in games with unusually high score on board. Earlier attempts from ^[8]Shah, Hazarika and Hazarikashowed a way to analyze players using Factor analysis, teams do spend a lot of time, discussing and making note on conditions, potential line ups of their opponents and possible winning strategy before the start of the match.

While [9] Lemmer (2008a) suggested that batsman is more likely to double his score if he was not dismissed, but given the nature of this particular format, gaining wickets is not the only way to claim victory.

In other words, it is important to not only bag wickets of key players, but also ensure that other batsman do not get enough opportunities to score runs either. This will allow them to make more mistakes, and subsequently increasing probability to get their wickets too. Therefore, in this format of, number of wickets are not the only thing that matters, but also the impact of the wicket obtained counts here. Similarly, economy of a bowler is not the only true parameter to judge the efficiency of one. Economy of two hypothetical bowlers with identical economy is being shown in the table below. However, a breakdown into more minute and detailed parameters is done to analyze the difference between their characteristics. In the given table below, both Bowler 1 and Bowler 2 have same average, they both are equal economically, if going with an assumption that this event after numerous repetition yield same result. Given that Bowler 1 concedes runs on his last bowl, in comparison to Bowler 2, this gives Bowler 1 an advantage over Bowler 2, as he is more likely to deliver an over without conceding any runs, in comparison to Bowler 2.

This is pivotal in realizing the point. If table is reviewed again, the distribution of runs across the ball delivered is more scattered if the deviation is being considered as ideal.

Table 1. Possible combinations of distributions of runs in an over.

	Bowler-1	Deviation	Bowler -2	Deviation
First Ball	0 runs	-1	1 runs	0
Second Ball	0 runs	-1	1 runs	0
Third Ball	0 runs	-1	1 run	0
Fourth Ball	0 runs	-1	1 run	0
Fifth Ball	0 runs	-1	1 run	0
Sixth Ball	6 runs		1 run	0

Parameters of selections

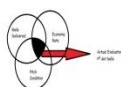


Fig-1: Factors required for Evaluation of dot balls

A. Presumption

Before laying out the mathematical formulation on a database, it is important to lay down some pre-requisite definitions and assumptions made for computing certain values. While many factors govern the performance of a bowler, as noted by [10] Bailey and Clarke. However, this paper aims to find a mathematical relation in order to find the impact of a bowler. Bearing that in mind, the parameters are narrowed to namely economy, number of wickets, and dot balls.

Another quantity named as [11] X-Factor (X) is also added in the formula. This is a product of strike rate of a bowler and Economy of a bowler. Where, strike rate denotes as “number of deliveries required to get a wicket”.

However, it might be noted that strike rate and total number of wickets, as mentioned above, are identical in properties, yet there is inclusion of such absolute and an aggregate quantity.

The reason behind such decision is that while form of a player is subjected to a change, but his absolute achievements cannot be ignored. That means, while strike rate can be a measure for temporary or “current” performance of a bowler, but number of wickets taken is a constant reminder about his level as a performer

B. Selection of Parameters

Following four parameters generally are preferred for selection of best bowler in a team.

P1: Number of dot balls per match.

P2: Number of wickets per match.

P3: Economy rate per over.

In limited over format of Cricket, **P3** becomes a prime for selecting a bowler in a team. However, in most of cases **P1** remains ignored. Assuming a situation where a bowler throws 50% of his delivered balls as dot balls, then irrespective of conceding huge number of runs. The selection of such bowler is questionable in playing XI. This paper considers the importance of this particular parameter **P1** for selecting a bowler into a playing eleven.

Like mentioned earlier, this format delivers great amount of runs in limited overs, thus if bowlers in an ideal situation, manage to scoop all of their delivered balls as dot balls, it is much more profitable in contrast to bowler poaching more wickets but conceding more runs. Dot balls add up to bowler’s credentials especially when the batting team is consistently scoring runs.

Dot balls also improves economy **P3** as well. If a team of T-20 cricket has five bowlers have high percentage of dot balls delivered, then the probability of restricting the batting team to a lesser total. This, to an extent, increases the chance of winning of a team with more dot balls than their opponent does that day.

Mathematically, $P1 = F^{-1}(P3)$, that is **P3** is inversely proportional to **P1**. However, **P1** is directly proportional to **P2**.

Combining both, it can write $P1 \propto (P2/P3)$.

Finally, $P1 = k P2/P3$, where k is a constant

(1)

So now, it is possible to frame an expression to explore the selection process in order to pick a bowler. Let

Q1: Dot Ball Percentage (Number of dot balls/Total Number of Balls).

Q2: Number of Wickets Taken

Q3: X-factor of the bowler

Where X-Factor is (Number of Balls/Wickets Taken)*(Runs Conceded/Number of overs)

Since twenty over format is the shortest format in cricket, the amount of runs being scored in lesser overs is very high. It is often a case where teams scoring huge totals with run-rate suggesting that of a figure which is not attained easily in other formats. In such cases, generating situations where dot balls can be thrown more is vital. Events where a batsman fails to score runs in a delivery certainly adds pressure on them to score more aggressively on the following delivery. So if teams focus on earning more dot balls, this will force their opponents to take bolder steps, just prompting more mistakes to capitalize upon. This approach combined with tactics supporting it can possibly reap rewards. As the results will be discussed further, a formula can now be created upon which bowlers can be judged to support this formula.

Bowling Score = $[0.6 \times (Q1/10)] + (Q2 \times 0.35) - [0.2 \times (Q3/100)]$ (2)

The selecting criteria maybe as below-

- If Bowling Score is greater than 3.75 the player should be selected in playing eleven with 100% probability.
- If Bowling Score belongs to $3.75 > \text{Bowling Score} > 3$, inclusion of player should be under consideration.
- If Bowling Score is under 3, then player can be ignored and team should look for other options at their disposal.

Now to implement the formula, given data of some of the best performers of Indian Premier League (IPL) of 2016 edition on basis of early performances.

Table-2: Evaluation table of best bowlers and decision regarding their inclusion in the team

C. Selection of bowlers on a given scenarios

Now with help of Bowling Score, X-Factor and dot ball percentage, IPL teams' performance can now be analyzed in more detail; it is possible to point out where a losing team were outpaced by their opponent. Such tabulations are enormously important in shaping strategy for upcoming matches and also ironing out issues that are effecting their chances of winning matches in the past, as shown below. It is interesting to observe that bar one match, all the teams with higher dot ball percentage emerged victorious. This further solidifies the theory that **"teams should focus on earning more dot balls to restrict run scoring opportunities"** which this paper presents.

And as for the one anomaly in this database: the match where a team lost despite having higher dot ball percentage, it is important to know that the losing team (Kolkata Knight Riders in this case) had a higher X-Factor than the winners, Mumbai Indians. Therefore, while a high dot ball

percentage provides an almost absolute chances to gain victory, a healthy tally of X-Factor is also important in

MuruganAshwin	76	83	26	46.67	6	4.72496	Yes
Dhawal Kulkarni	24	19	16	66.67	2	4.5862	Yes
Sandeep Sharma	55	50	28	50.9	4	4.301222	Yes
Amit Mishra	50	48	22	44	5	4.19	Yes
Mitchell McClean	66	73	29	36.8	6	4.162	Yes
Andre Russell	68	111	27	39.7	5	3.85754	Yes
RavindraJadeja	49	48	20	40.81	4	3.7016	Yes
Dwayne Bravo	74	104	21	28.37	6	3.58842	Yes

maintaining a discipline over bowlers and thus, encouraging better results. This is evident in the table-3, where the two best sides in Kolkata Knight Riders and Delhi Daredevils, with highest X-Factor and Dot Ball percentage had the better record after round three (2 wins and 1 defeat for both of them). While, a slightly lower X-Factor could have given Kolkata a perfect record of 3 wins, their incredible show of disciplined bowling have still maintained a strong hold over other IPL teams. This is result of some shrewd team selection and bowling instructions, and it clearly shows in statistics too.

Table-3: Comparative analysis of teams during IPL-2015

	Total Balls	Runs Conceded	Dot Balls	Dot Ball Percentage	Wickets Taken	Bowling Score	Decision of Selection in team
	Total Balls	Runs	Dot Balls	Dot Ball %	Wickets	X-Factor	Result
RPS	125	121	66	52.8	8	94.53125	Lost
MI	92	126	38	41.3	1	805	Won
KKR	110	98	66	60	10	61.954023	Won
DD	90	99	42	46.67	1	90	Lost
GL	124	161	39	31.45	6	166.36667	Won
KxiP	109	162	31	28.44	4	377.06897	Lost
SH	125	227	35	28	4	354.6875	Lost
RCB	125	182	46	36.8	6	189.58333	Won
MI	127	187	25	19.68	5	237.49	Won
KKR	119	188	40	33.61	4	286.82051	Lost
GL	120	163	38	31.67	5	195.6	Won
RPS	116	164	35	30.17	3	330.27778	Lost
DD	123	111	68	55.28	7	97.521429	Won
KxiP	84	113	31	36.9	2	356.84211	Lost
KKR	125	139	49	39.2	6	144.79167	Won
SH	115	146	35	30.43	2	585.60439	Lost
GL	127	143	56	44.09	8	113.50625	Won
MI	124	147	43	34.67	7	372	Lost
KxiP	123	152	41	33.34	6	155.8	Won
RPS	114	153	36	31.57	4	236.9837	Lost

Table-4: Evaluation table of teams under terms derived in paper

	Total No. of Balls	Total Runs Conceded	Dot Balls %	Wickets	X-Factor
RPS	355	438	38.591549	15	563.36957
MI	343	460	30.90379	12	241.6973
KKR	354	425	43.785311	20	133.37766
DD	334	401	51.643192	12	206.30622
GL	371	467	35.849057	19	151.97982
KxiP	316	427	32.911392	12	221.78172
SH	240	373	29.166667	6	390.57592
RCB	246	374	31.707317	9	261.44928

However, to achieve similar success on longer stretch of games, teams need to analyze other external factors which exist outside the statistical matrix. Smaller details like nature of pitch, temperature, moisture, wind speed, the grass on the ground & pitch; all of them are necessary to formulate a strategy. This will exactly determine the requirement of number of bowlers and exactly the type of bowlers team needs. And with the given data available, it is easier to evaluate not the parameters defined in this paper, but also dissect and observe the popular parameters, namely Economy.

As mentioned earlier, Economy is not a perfect indicator of efficiency of a bowler. However, if Economy itself is divided into a sub-atomic level of it own, then it becomes a clearer parameter. With Economy now divided into **Initial, Middle and Final**

E.Over Economy, it is possible to study the nature about the bowler. This is useful to generate a game plan where the captain of the team is aware about previous performance of a bowler in certain overs of the game.

Thus, captain can potentially utilize the bowler's ability to his team's benefits as shown in Table-4.

For programming purposes, this is where Machine Learning comes into play. With the help of Decision Tree, computer can identify the need of a particular bowler for all types of situations and make a decision to select a bowler that is capable of delivering more dot balls or taking wickets in next over. This will maximize the chance of winning for a team, taking accounts of all natural and sporting factors of the game,

Table-5: Evaluation table of bowlers and differentiation of economy

	Total Balls	Runs Conceded	Dot Balls	Dot Ball	Wickets	Bowling Score	Initial Over Economy	Middle	Final Over Eco.
				Percentage	Taken			Over Eco.	
MuruganAshwin	76	83	26	46.67	6	4.72496	20	5.34	6.8
Dhawal Kulkarni	24	19	16	66.67	2	4.5862	3.67	8	-
Sandeep Sharma	55	50	28	50.9	4	4.301222	5.14	7.5	-
Amit Mishra	50	48	22	44	5	4.19	-	4.4	-
Mitchell McCleanaghan	66	73	29	36.8	6	4.162	7.67	6	7
Andre Russell	68	111	27	39.7	5	3.85754	8.2	4.67	13.67
RavindraJadeja	49	48	20	40.81	4	3.7016	-	7.5	2
Dwayne Bravo	74	104	21	28.37	6	3.58842	16	5	9.125

D.Projection

While pitch conditions are taken into account in previous section. This allows us to have an access to an immense database where not only there can be evaluation of a better bowler. But also find out how and what conditions bring the best out of him. This intelligence enables teams to make the best possible combination of bowlers to enhance their chances of winning on the basis of their statistical performances. Using mathematical operation of projection, teams will be observe the progress the player is making and predict his possible output in near future.

Such analysis is pivotal in shaping teams and in case of Indian Premier League, providing a chance to invest on

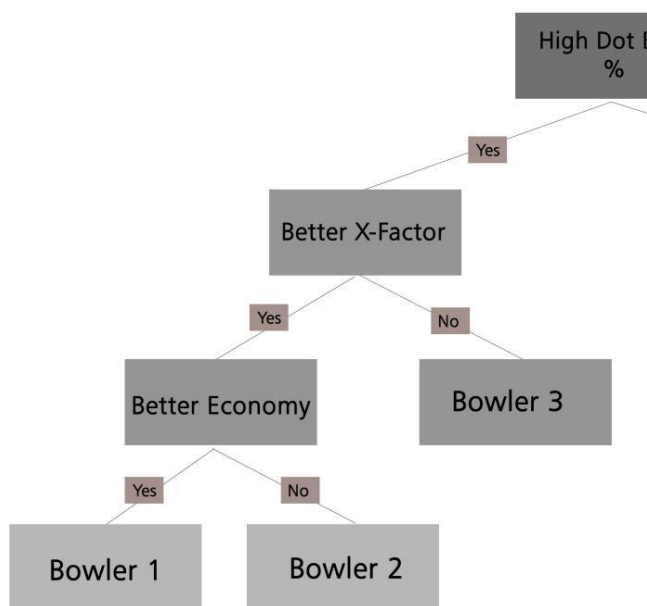
players who can strengthen the team to follow a particular model. Team can effectively change their strategy ahead of the game. In case of facing a team superior on a pitch where bowlers face a hard time, targeting on gaining more dot balls can change game into their favor. Similarly, up

against a team not up to the standards, using bowlers more likely to get a wicket might come in handy.

While teams in real life, do go for later approach, they tend to restrict teams' boundaries rather than blocking any opportunity to score a run. Dot balls do increase the pressure on batsman. And results are discussed in the next section.

Involvement of Computing to improve decision making in Cricket

Figure-2: Decision Tree to select a bowler.



V.CONCLUSION

Early results are encouraging to say the least. With all the accounts of eleven matches computed, all but one match were won by teams with higher dot ball percentage. Similarly, on the basis of individual performances, the initial performances of the bowlers with better **Bowling Score** went on to perform exceedingly well in 2016 edition of Indian Premier League. Mustafizur Rahman, was one such revelation of that season. And his statistics do suggest that he has a potential to be a great spinner to target dot balls as well as wickets.

So while data sample is small, but patterns are beginning to emerge. Therefore, this idea is further extended to a full length program to trace such tiny statistical details and translate them into desired results.

However, it is important to note that coefficients and constants present are under consideration and will be evaluated and manipulated if required in future progress of this paper. Not just coefficients and constants, but whole mathematical morals will be enhanced accordingly as per more matches come in and in depth analysis of them.

Furthermore, weight-age is on presumption basis and will be walked upon in due course of time.

This paper has suggested inter relationship among three variables which are economy rate, wicket taken, dot ball delivered. An indicator based formula has been suggested for choosing a bowler into the playing eleven when the past performance has been taken into consideration. The main focus in developing indicator is high preference to dot balls delivered rather than wicket taken and economy rate. It has been discussed that since these are inter related, so focusing more on dot ball delivery will lead to automatic incorporation of other parameters. It has been suggested that 60% preference should go to dot ball delivered while developing an indicator.

REFERENCES

1. Batdelger Doljin*, Franz Konstantin Fuss. Development of a smart cricket ball for advanced performance analysis of bowling.
2. Justham L., West A. (2006) The Use of System Analysis and Design Methodology in the Development of a Novel Cricket Bowling System. In: Moritz E.F., Haake S. (eds) The Engineering of Sport 6. Springer, New York, NY
3. Ian Renshaw and Malcolm M. Fairweather, Cricket bowling deliveries and the discrimination ability of professional and amateur batters. Journal of Sports Science Volume 18, 2000- Issue 12.
4. A.F. Burnett, B.C. Elliott and R.N. Marshall, The effect of 12 over Spell on Fast Bowling Technique in Cricket. Journal of Sports Science, 1995, 13, 329-341.
5. Craig A. Ranson, Angus F. Burnett, Mark King, Nitin Patel, and Peter B. O'Sullivan, The relationship between bowling action classification and three-dimensional lower trunk motion in fast bowlers in cricket. Journal of Sports Sciences 26, 267-276, 2008.
6. Aaron Chin, The off-break and "doosra": Kinematic variations of elite and sub-elite bowlers in creating ball spin in cricket bowling. Sports Biomechanics Volume 8, 2009- Issue 3.
7. Hernamus Lemmer, South African Journal for Research in Sport, Physical Education and Recreation - A measure for the batting performance of cricket players. South African Journal for Research in Sport, Physical Education and Recreation, Volume 26, Issue 1, Jan 2004, p. 55 - 64
8. Sricharan Shah, Partha Jyoti Hazarika, Jiten Hazarika. A Study on Performance of Cricket Players using Factor Analysis approach. Vol 8, No. 3,
9. International Journal of Advanced Research in Computer Science.
10. LEMMER, H.H. (2008a). Measures of batting performance in a short series of cricket matches. South African Statistical Journal, 42(1): 83-105.
11. Bailey, M.J. & Clarke, S.R: Predicting the match outcome in one day international cricket matches, while the match is in progress. Journal of Science and Sports Medicine, 5, 480-487 (2006).