A DEA Model for Selection of Indian Cricket Team Players

Riju Chaudhary[1], Sahil Bhardwaj[2], Sakshi Lakra[3]

[1]Department of Mathematics, University of Delhi, Delhi
[2,3]Amity Institute of Applied Sciences, Amity University, Noida
Amity Institute of Applied Sciences, Amity University, Noida
[1]riju.chaudhary@gmail.com, [2]sahilbhardwaj3@yahoo.com, [3]sakshilakra1996@gmail.com

Abstract: Cricket has become one of the most popular sport tournaments in the world over the last few decades. To form a winning team for cricket tournaments,analysis &evaluation of past performance of the cricket players is required to select players strategically. This paper suggests an effective technique for selecting Indian Cricket Team members by measuring the efficiency of cricket players using Data Envelopment Analysis (DEA). The results obtained by the technique suggested in this paper are in sync with the team players selected by the cricket board.

Keywords – Data Envelopment Analysis, Efficiency Evaluation, Cricket, Test Match, Player Selection

I. INTRODUCTION

Cricket is one of the most popular game played worldwide. Many studies have been conducted to analyse various dimensions of the game such as batting strategies, bowling strategies, efficiency of players. Winning a game of cricket depends a lot on the team players selected. This paper suggests a technique called as **Data Envelopment Analysis (DEA)**, to calculate the efficiency of players based on their past performances and selects the best ones to form a winning team. There has been a tremendous growth to the applications of DEA in many areas, which involve the use of multiple inputs to yield multiple outputs. DEA assesses the relative efficiency of a set of comparable **Decision Making Units (DMUs)** by using a ratio of the weighted sum of outputs to the weighted sum of inputs.

Preston, I., & Thomas, J [10] aimed at identifying batting strategies employed in the game of limited overs. Bracewell and Ruggiero [4] developed control charts for assessing the performance of batsmen over a period of four years. Barr, G.D. I., Holdsworth, G. C., & Kantor, B.S [3] evaluated performance of batsmen and bowlers of Australian team in the Cricket World Cup, 2007. Application of neural networks for making valuable decision in the process of team selection was done by Iyer and Sharda [9]. Amin, G.R. and Sharma, S.K [1] suggested an approach for finding the best efficient players for which they took the data of the Indian Premier League (IPL) 2007. Amin, G.R., and Sharma [2] proposed a two stage method for evaluating the performance of Indian batsmen using Ordered Weighted Averaging (OWA) based on data set taken from Indian Premier League (IPL) 2011 and ranked them using regression. Singh and Adhikari [11] incorporated available bowling resources and individual's excellence obtained with the help of Data Envelopment Analysis (DEA).

Studies suggest that at times players do not show their actual performances due to unethical practices such as match fixing. Also, not every player gets a chance to perform in T20 matches, but every player gets a chance to perform in Test Matches where they tend to show their actual performances. Hence, the data set taken from a test match or even a One Day International (ODI) match can be considered as a more appropriate source to obtain reliable efficiency & precision. In this paper scores of test matches of Indian Cricket team has been taken from the official website of ESPNcricinfo [8]. The efficiency of each batsmen & each bowler is thus calculated, players are chosen according to the requirement and their efficiency scores to form a team. On verifying our results, it is found that the team proposed by the model is in-sync with the team selected by the Board of Control for Cricket in India (BCCI).

II. DEA MODELS

The objective of DEA model, suggested by Charnes& Cooper in 1978 [5], is to calculate the efficiency of aDecision Making Unit (DMU). Efficiency is obtained by maximizing the ratio of weighted outputs to weighted inputs of the target DMU, DMU_0 relative to the efficiency scores of other DMUs into consideration. Efficiency lies between 0 and 1.

A. Basic DEA Model

$$\max \mu_0 = \frac{\sum_{r=1}^{s} \alpha_r x_{r0}}{\sum_{i=1}^{m} \beta_i y_{i0}}$$

$$s.t. \quad \frac{\sum_{r=1}^{s} \alpha_r x_{rj}}{\sum_{i=1}^{m} \beta_i y_{ij}} \leq 1 \quad j = 1, 2, \dots, n$$

$$\alpha_r \geq 0 \quad r = 1, 2, \dots, s$$

$$\beta_i \geq 0 \qquad i = 1, 2, \dots, m$$

The model determines the efficiency of n DMUs where x_{1j}, x_{2j}, …., x_{mj} are the m inputs and y_{1j}, y_{2j},….,$y_{sj}$are the s outputs of the j^{th} DMU and α_r, $\beta_i \geq 0$ are the weight vectors associated

978-1-5386-9346-9/19/$31.00 ©2019 IEEE

with r^{th} output and i^{th} input of DMU_0, respectively to be determined.

A. Fractional to Linear Model

To obtain the solution of a fractional programming problem, it needs to be first converted into a linear programming problem using a method given by Charnes and Cooper. Since the basic DEA model involve fractions, its corresponding linear programming problem is mathematically expressed as:

$$\max_{u_r, v_i} h_0 = \sum_{r=1}^{s} u_r y_{r0}$$

$$s.t. \quad \sum_{i=1}^{m} v_i x_{i0} = 1$$

$$\sum_{r=1}^{s} u_r y_{rj} \leq \sum_{i=1}^{m} v_i x_{ij} \; j = 1,2,\ldots.,n$$

$$u_r \geq 0 \qquad r = 1,2,\ldots.,s$$

$$v_i \geq 0 \qquad i = 1,2,\ldots.,m$$

where,

$$u_r = t\mu_r (r = 1,2,\ldots\ldots,s),$$

$$v_i = t\gamma_i (i = 1,2,\ldots\ldots,m) and\ t = \left(\sum_{i=1}^{m} \gamma_i x_{i0}\right)^{-1}$$

The dual for the above model can be expressed as follows:

$$min\ \theta_0$$

$$s.t. \sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta_0 x_{i0} \qquad i = 1,2,\ldots\ldots,m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{r0} \qquad r = 1,2,\ldots\ldots,s$$

$$\lambda_j \geq 0 \qquad j = 1,2,\ldots\ldots,n$$

A. DEA Model with Explicit Output

For selection of members of a cricket team, it is assumedthat there are 'n' players. The 's'outputs of j^{th} player are denoted by the set $y_j = (y_{1j}, y_{2j}, \ldots\ldots, y_{sj})$. A general DEA model, with such explicit outputs, for calculating the efficiency of $thek^{th}$ player (k = 1,2,...,n) is as follows :

$$max\ h$$

$$s.t. \sum_{j=1}^{n} y_{rj}\lambda_j - y_{rk} h \geq 0 \qquad r = 1,2,\ldots.,s$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \geq 0 \qquad j = 1,2,\ldots\ldots,n$$

The dual of above model can be written as:

$$min\ u_0$$

$$s.t. \ -\sum_{r=1}^{s} y_{rj} u_r + u_0 \geq 0 \qquad j = 1,2,\ldots\ldots\ldots,n$$

$$\sum_{r=1}^{s} y_{rk} u_r = 1$$

$$u_r \geq 0 \qquad r = 1,2,\ldots\ldots\ldots,s$$

III. EFFICIENCY EVALUATION OF BATSMEN

- Runs : Total number of runs scored by a player while batting in cricket matches. This factor is very important for measuring the efficiency of a player.
- Strike Rate (SR): Ratio of total runs scored to the total bowls faced multiplied by 100. It simply means the number of runs scored by a player per 100 balls.

  Let the strike rate be α, then

  $$\alpha = \frac{\text{Total runs scored by a player}}{\text{Total number of balls faced by a player}} * 100$$

- Batting Average (Ave):A batting average represents how many runs, on average, a batsman scores before getting out. The higher the batting average, the better the batsman's ability to score runs without getting out.

  Batting Average:Runs divided by (number of times out) = Runs divided by (Innings - Not Outs)

  $$\frac{\text{Total runs scored by a player in all innings}}{\text{Total innings playes} - \text{Not out innings}}$$

  $$\frac{\text{Total runs by a player}}{\text{Total number of times a player out}}$$

- 100s - Total number of 100's hits by a batsman i.e. number of full centuries hit by a batsman.

- 50s - Total number of 50's hit by a batsman i.e. number of half centuries hit by a batsman.
- 4s& 6s - Total number of boundaries hit by a batsman for 4& 6runs.
- 6s - Total number of boundaries hit by a batsman for 6 runs.

Batting statistics of each player and their calculated efficiency scores are tabulated in Table I.

TABLE I: BATTING STATISTICS OF INDIAN CRICKET PLAYERS

Players	Runs	Ave	SR	100s	50s	4s	6s	DEA Score
M Vijay	3933	39.33	46.56	12	15	463	33	1
V Kohli	6331	54.57	58.26	24	19	700	18	1
CA Pujara	4905	49.54	47.05	15	19	588	9	1
PP Shaw	237	118.5	94.04	1	1	34	1	1
RA Jadeja	1395	32.44	64.1	1	9	140	39	1
AM Rahane	3271	41.4	50.91	9	15	371	23	0.912117
S Dhawan	2315	40.61	66.94	7	5	316	12	0.902682
HH Pandya	532	31.29	73.88	1	4	68	12	0.891401
RR Pant	346	43.25	73.61	1	2	39	12	0.877216
KL Rahul	1848	37.71	58.11	5	11	220	13	0.845151
R Ashwin	2331	29.5	55.31	4	11	265	14	0.820802
KD Karthik	1025	25	49.27	1	7	134	4	0.657981
GH Vihari	56	28	43.07	0	1	7	1	0.469358

Using the above data, following equations can be formulated to determine the efficiency of M.Vijay:

Min (β)

Subject to Constraints:-

$3933\lambda_{vijay} + 6331\lambda_{kohli} + \ldots\ldots\ldots + 1025\lambda_{kartik} + 56\lambda_{vihari} >= 3933\beta$

$39.33\lambda_{vijay} + 54.57\lambda_{kohli} + \ldots\ldots\ldots + 25\lambda_{kartik} + 28\lambda_{vihari} >= 39.33\beta$

$46.56\lambda_{vijay} + 58.26\lambda_{kohli} + \ldots\ldots\ldots + 49.27\lambda_{kartik} + 43.07\lambda_{vihari} >= 46.56\beta$

$12\lambda_{vijay} + 24\lambda_{kohli} + \ldots\ldots\ldots + \lambda_{kartik} >= 12\beta$

$15\lambda_{vijay} + 19\lambda_{kohli} + \ldots\ldots\ldots + 7\lambda_{kartik} + \lambda_{vihari} >= 15\beta$

$463\lambda_{vijay} + 700\lambda_{kohli} + \ldots\ldots\ldots + 134\lambda_{kartik} + 7\lambda_{vihari} >= 463\beta$

$33\lambda_{vijay} + 18\lambda_{kohli} + \ldots\ldots\ldots + 4\lambda_{kartik} + \lambda_{vihari} >= 33\beta$

$\lambda_{vijay} + \lambda_{kohli} + \ldots\ldots\ldots + \lambda_{kartik} + \lambda_{vihari} = 1$

$\lambda_{vijay}, \lambda_{kohli}, \ldots\ldots\ldots \lambda_{kartik}, \lambda_{vihari} >= 0$

Similarly efficiency of every other batsmen can be determined.

IV. EFFICIENCY EVALUATION OF BOWLERS

- Overs - One over is equal to delivery of 6 balls by a bowler. Here Over means total number of overs delivered by a bowler.
- Maiden Overs (Mdns) - Maiden over is an over in which no runs have been scored by a batsman.
- Runs - Number of runs given by a bowler during his bowling turn.
- Wickets (Wkts) - Number of wickets taken by abowler.
- Average (Ave) - Number of runs per wicket taken.
  Let bowling average be α

  Then, $\alpha = \frac{\text{Total runs}}{\text{Wickets taken}}$
- Economy Rate (Econ) - Bowling economy rate is defined as total number of runs conceded divided by number of overs delivered.

  $$\text{Economy rate} = \frac{\text{Total runs conceded}}{\text{Numver of overs delivered}}$$
- Strike Rate (SR) - Bowling Strike Rate is a measurement of a bowler's average number of balls bowled for every wicket taken.

  $$\text{Strike rate} = \frac{\text{Total number of bowls bowled}}{\text{Total number of wickets taken}}$$
- 5's & 10's - This parameter is only used for test matches means 5 or 10 wickets in a single match this parameter is very effective for judging a bowler for test match.

Bowling statistics of each player and their calculated efficiency scores are tabulated in Table II.

V. RESULT

Efficiency scores of Indian cricket players for selection of a potentially winning team for test match serieshas been evaluated. Expert's advices play an important role while selecting a team for test match. According to experts every player has to do batting, therefore, a bowler having better batting strike rate is given more preference over other players. As per the efficiency scores, our selected players are:

V Kohli, CA Pujara, PP Shaw, M Vijay, AM Rahane, S Dhawan, RA Jadeja, R Ashwin, I sharma, GH Vihari, UT Yadav, H Pandaya, M Shami.

TABLE II: BOWLING STATISTICS OF INDIAN CRICKET PLAYERS

Players	Overs	Mdns	Runs	Wkts	Ave	Econ	SR	5	10	DEA Score
R Ashwin	2975.1	607	8551	336	25.44	2.87	53.1	26	7	1
KuldeepYadav	133	16	480	19	25.26	3.6	42	1	0	1
I Sharma	2763	540	8893	256	34.73	3.21	64.7	8	1	1
GH Vihari	10.3	1	38	3	12.66	3.61	21	0	0	1
UT Yadav	1072.5	183	3844	117	32.85	3.58	55	2	1	1
HH Pandya	156.1	19	528	17	31.05	3.38	55.1	1	0	0.965447
Mohammed Shami	1123.4	191	3835	128	29.96	3.41	52.6	3	0	0.964748
RA Jadeja	1824.2	462	4348	185	23.5	2.38	59.1	9	1	0.929057
B Kumar	558	141	1644	63	26.09	2.94	53.1	4	0	0.874516
JJ Bumrah	245.3	47	716	28	25.57	2.91	52.6	2	0	0.865876

Indian cricket team selected by BCCI for India - England test match series(2018) include the following players:

V Kohli, CA Pujara, PP Shaw, AM Rahane, S Dhawan, KL Rahul, K Nair, R Ashwin, RA Jadeja, I Sharma, GH Vihari, UT Yadav, H Pandaya, M Shami, S Thakur, J Bumrah, RR Pant(Wicket Keeper).

VI. CONCLUSION

A DEA model was formulated incorporating batting-bowling parameters of cricket players. Efficiency of each player served as a basis for his inclusion in the team. On comparing the players selected by BCCI for India – England test series with the players selected through our technique matches to a great extent. The players which were efficient according to model proposed in the paper were actually the players who played from Indian Cricket team against England. This shows that the proposed model can be used for selecting a cricket team.

REFRENCES

[1] Amin, G.R. and Sharma, S.K. 2014. Cricket team selection using data envelopment analysis. European Journal of Sport Science, 14(1), pp. 369-376.

[2] Amin, G. R., & Sharma, S. K. (2014). Measuring batting parameters in cricket: A two-stage regression-OWA method. Measurement, 53, 56-61.

[3] Barr, G.D. I., Holdsworth, G. C., & Kantor, B. S. (2008) Evaluating performance at the 2007 cricket world cup. South African Statistical Journal, 42, 125-142.

[4] Bracewell, P. J., & Ruggiero, K. (2009). A parametric control chart for monitoring individual batting performances in cricket. Journal of Quantitative Analysis in Sports, 5, 1_19.

[5] Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. European journal of operational research, 2(6), 429-444.

[6] Cooper, W. W., Seiford, L. M., & Tone, K. (2007). Data envelopment analysis: a comprehensive text with models, applications, references and DEA-solver software. Springer Science & Business Media.

[7] Cook, W. D., &Seiford, L. M. (2009). Data envelopment analysis (DEA)–Thirty years on. European Journal of Operational Research, 192(1), 1-17.

[8] Cricket Teams, scores, Stats, News, Fixtures, Results, Tables. ESPNcricinfo, www.espncricinfo.com.

[9] Iyer, S. R., &Sharda, R. (2009). Prediction of athletes performance using neural networks: An application in cricket team selection. Expert Systems with Applications, 36, 5510_5522.

[10] Preston, I., & Thomas, J. (2000). Batting strategy in limited overs cricket. Journal of the Royal Statistical Society, Series D (The Statistician), 49, 95-106.

[11] Singh, S., &Adhikari, A. (2015). A new net resource factor based alternative method to calculate revised target in interrupted one day cricket matches. American Journal of Operations Research, 5(03), 151.