South African Journal for Research in Sport, Physical Education and Recreation, 2012, 34(2): 95-103.
Suid-Afrikaanse Tydskrif vir Navorsing in Sport, Liggaamlike Opvoedkunde en Ontspanning, 2012, 34(2): 95-103.
ISBN: 0379-9069

INDIVIDUAL MATCH APPROACH TO BOWLING PERFORMANCE MEASURES IN CRICKET

Hermanus H. LEMMER
Department of Statistics, University of Johannesburg, Johannesburg, Republic of South Africa

ABSTRACT

Match conditions can play a significant role in player performances in a cricket match. If the pitch is in a good condition, the batsmen can achieve good scores, making it difficult for the bowlers. In the case of an uneven pitch or adverse weather conditions, the bowlers may have the upper hand. In order to measure bowlers' performances in a fair way, it is desirable that the match conditions should be taken into account in the calculation of bowling performance measures. The purpose of this study was to provide a fair method to select the best bowler of a series or tournament. This is achieved by means of the individual match approach, which is introduced in the present study. Its essence is to scale the number of runs conceded by a bowler down when the batting conditions are favourable and to scale them up when batting conditions are adverse. The measure proposed is used to rank the bowlers in the ICC World Twenty20 series of 2010. The use of these adjusted runs can have a significant effect on the ranking of bowlers compared to the situation where the original number of runs is used in the same measure.

Key words: Bowling performance; Cricket; Cricket pitch; Wicket weights.

INTRODUCTION

During the past decade or two, a large number of papers have been published on cricket performance measures and prediction methods. The majority of these papers concentrate on batting performances in limited overs matches. The most important are those dealing with the Duckworth-Lewis method, which is used to adjust the target in the case of rain-interrupted matches (Duckworth & Lewis, 2002; O'Riley & Evans, 2006). Much less attention has been paid to bowling measures.

The traditional bowling performance measures are the bowling average, the economy rate and the strike rate (Kimber, 1993). Various authors used graphic methods based on these measures to compare bowlers' performances (Kimber, 1993; Skinner, 1995; Barr & Kantor, 2004; Barr *et al*., 2008; Van Staden, 2009). The approaches are very similar, with the strike rate being plotted on the vertical axis and the economy rate on the horizontal axis. The bowlers whose co-ordinates lie closest to the origin are seen to be the best ones. The definition of the strike rate used by Barr and Kantor (2004) and Barr *et al*. (2008) is the inverse of the definition used by the others. Measures, which combine two of the measures, can be found in Croucher (2000), Barr and Kantor (2004) and Basevi and Binoy (2007).

Lemmer (2002) proposed a measure based on all three measures and gave a further refinement in Lemmer (2004). Other approaches can be found in Bairam *et al*. (1990), who used a production function approach to determine the best batting and bowling strategy to maximize the probability of winning. Cohen (2002) used the strike rates of bowlers to calculate the probability of dismissing the opposing team. Beaudoin and Swartz (2003) proposed a statistic 'runs per match' that can be used for batsmen and bowlers alike. It utilises the Duckworth/Lewis resource table in a very sensible way. Its use is restricted, especially in the case of bowlers, by the fact that data recording is such that it requires an enormous effort to extract the necessary data for analyses. In the present study the method of Lemmer (2002) is adapted to the situation where a small number of matches had been played and the fact that match conditions might have differed and had to be taken into account.

BOWLING PERFORMANCE MEASURES

Traditionally the bowling average is defined by A=R/W, with R the number of runs conceded and W the number of wickets taken by the bowler, the economy rate by E=R/O, where O denotes the number of overs bowled, and the strike rate by S=B/W, with B the number of balls bowled. These definitions are in agreement with those shown in television broadcasts and used on the Cricinfo (2010b) and other cricket web pages. Bairam *et al*. (1990) called the strike rate 'attacking bowling' and the economy rate 'defensive bowling'. In the graphic methods some of the measures are defined differently. Barr and Kantor (2004) defined the strike rate as SB=W/B, the economy rate as EB=R/B and the average as AB=W/R. It is easy to see that AB=1/A, EB=E/6 and SB=1/S in terms of the traditional definitions. See also Barr *et al*. (2008).

According to Kimber (1993) the average has traditionally been used to compare bowlers, but the economy rate and strike rate have more recently increased in popularity. He mentioned that as early as 1950, Sir Donald Bradman used the strike rate as a measure of the 'worth' of a bowler (Bradman, 1950). Each of these measures is important in its own right, but some authors use combinations of these measures. Croucher (2000) defined the bowling index BI=A×S, and used this to rank bowlers. Basevi and Binoy (2007) used CALC=A×E/6. Barr *et al*. (2008) used as the measure of bowling performance a weighted product of their strike rate and average, namely:

$$\text{BHK} = \text{SB}^{\alpha} \times (\text{AB})^{1-\alpha}, \; 0 \leq \alpha \leq 1.$$

By using different values of α, the importance of the strike rate relative to the average can be varied. They used α=0.5 and α=0.75 for illustrative purposes. A more comprehensive measure that has been designed to take A, E and S deliberately into account is the combined bowling rate

$$\text{CBR} = 3\text{R}/(\text{W} + \text{O} + \text{W}\times\text{R}/\text{B})$$

for use in limited overs matches (Lemmer, 2002). In CBR the economy rate is heavily weighted because in a limited overs match it is much more important to limit the scoring rate of batsmen than to take wickets. In Lemmer (2002), CBR was defined as the harmonic mean between A, E and S:

$$\text{CBR} = 3/(1/\text{A} + 1/\text{E} + 1/\text{S}) = 3\text{R}/(\text{W} + \text{O} + \text{W}\times\text{R}/\text{B}).$$

Among the measures mentioned, only CBR contains all of A, E and S explicitly.

It is interesting to note that CBR can also be written as CBR=c×RG, where c is a constant and RG=(A_s + E_s + S_s)/3 the arithmetic mean between the standardised values of A, E and S. Here A_s =A/AM, where AM denotes the average of A from a given data set (E_s and S_s have similar definitions) (Lemmer, 2002). The value of the constant c depends on the data set used and c=7.65 for the specific data set. This form gives a further perspective on CBR, but is not suitable for calculation purposes.

Many factors influence the performance of a bowler, such as the pitch condition, the weather, light quality, the strength of the opposing batsmen, the fielding of his own team members, the judgements of the umpires, etc. In the long run, most of these factors tend to balance out, but in a short series of matches (a World Cup series) such factors should be taken into account as far as possible in order to get a fair comparison between bowlers. None of the measures mentioned make specific provision for the case of a small number of matches. Lemmer (2005) showed how the measure CBR could be adjusted to make provision for this case. The rationale of his approach was that a bowler who had taken the wickets of the four top-order batsmen had performed better than one who had taken the wickets of the four tail-enders. It was proposed that one should not simply count the number of wickets taken by a bowler, but that weights should be allocated to the wickets taken and that one should work with W*, the sum of the weights of the wickets taken by the bowler. The formula CBR then becomes

$$CBR^* = 3R/(W^* + O + W^* \times R/B).$$

In the present study, involving Twenty20 matches, the weights for ODIs in Table 1 of Lemmer (2005) (Table 1) will be used, because both Twenty20 and ODI matches are limited overs matches. As soon as sufficient international Twenty20 data become available, wicket weights will have to be calculated for these types of matches.

TABLE 1: **WEIGHTS OF WICKETS ACCORDING TO BATTING POSITION**

Batting position	Weight	Batting position	Weight
1	1.30	7	0.98
2	1.35	8	0.79
3	1.40	9	0.59
4	1.45	10	0.39
5	1.38	11	0.19
6	1.18	**Total**	11.00

INDIVIDUAL MATCH APPROACH

A method designed to eliminate the effect of factors like pitch and weather conditions will now be introduced. Assuming that a fixed number of overs can be bowled in a match and that 20 wickets can be taken, the above-mentioned conditions will best be reflected by the total number of runs conceded. Under favourable batting conditions bowlers will concede more runs than under difficult conditions. In measuring a bowler's overall performance after a series of matches it should be taken into account that match conditions could have differed markedly between matches. It is therefore necessary for overall bowling performance measures to be adapted to take this into account.

The approach in this paper is as follows: if two bowlers had bowled the same number of overs and taken the same number of wickets, the one who had conceded the least number of runs had performed best. If they had conceded the same number of runs, the one who had played under the better match conditions had performed best, because he could limit the number of runs scored despite favourable batting conditions. It remains to quantify the degree of difficulty of the match conditions. A bowler's economy rate (or the average number of runs conceded per ball bowled) gives a good indication of his ability to restrict scoring. Now calculate the runs conceded per ball bowled, RPB=R/B for the bowler, and also, for all bowlers in the whole match, RPBM=total number of runs scored divided by the total number of balls bowled. If a bowler's RPB is less than RPBM, he performed better than the average for the match, so it is logical to scale his number of runs conceded, R, down by using the ratio RPB/RPBM in a suitable way.

By using a downscaling similar to that of BP_{26} in Lemmer (2008a), his adjusted number of runs conceded is taken as $AR=R\times(RPB/RPBM)^{0.50}$. Thus, the match conditions adjusted number of runs conceded, AR, replaces each bowler's number of runs conceded. This is done for every match in which the bowler had played. Let $R^{\#}$ denote the sum of his adjusted runs, AR, in the series. This is used in the formula of CBR*, which then becomes

$$CBR^{\#} = 3R^{\#}/(W^* + O + W^*\times R^{\#}/B)$$

with O the total number of overs bowled by the bowler in the series.

RESULTS

The measure $CBR^{\#}$ has been used to compare bowlers' performances in the ICC Twenty20 series played during March and April 2010. The full scorecards obtained from Cricinfo (2010a) have been used to calculate the sum of the weights of all the wickets taken by each bowler, W*. The bowling figures, obtained from Cricinfo (2010b), have been used to calculate the RPBM values for each match and the AR values for all the bowlers in all the matches. For comparative purposes it was required that a bowler should have bowled at least 12 overs (the number of overs that a bowler is allowed to bowl in 3 matches). The bowlers' statistics are given in Table 2, where they have been ranked according to $CBR^{\#}$.

Langeveldt bowled extremely well, and took the wickets of mainly good batsmen, resulting in W*=12.87 for his 11 wickets. His economy rate was almost the same as that of the group, with the result that his $R^{\#}$=104.36 differed very little from R=104. Sammy's very good economy brought R=72 down to $R^{\#}$=64.36, his match conditions adjusted number of runs conceded. This caused CBR=8.72 to be replaced by the more appropriate measure $CBR^{\#}$=7.97. Keeping in mind that small values of these measures indicate good performances, his value of $CBR^{\#}$ indicates an 8.57% better performance than does CBR*. In the case of Tait the 'improvement' was 11.14%. Watson, on the other hand, was very uneconomical, with the result that his value of $CBR^{\#}$ (=30.89) was 14.36% worse than that of CBR* (=27.01).

TABLE 2: **RANKING OF BOWLERS ACCORDING TO CBR#**

Rank	Name	O	R	R#	W	W*	CBR*	CBR#	A	E	S
1	C Langeveldt	16.0	104	104.36	11	12.87	7.29	7.30	9.45	6.50	8.73
2	DJG Sammy	13.4	72	64.36	6	5.91	8.72	7.97	12.00	5.27	13.67
3	DP Nannes	26.0	183	176.46	14	17.94	8.45	8.24	13.07	7.04	11.14
4	GP Swann	22.0	144	136.66	10	13.33	8.66	8.34	14.40	6.55	13.20
5	SW Tait	23.4	131	113.27	9	8.7	9.73	8.64	14.56	5.53	15.78
6	MG Johnson	22.2	145	133.45	10	11.57	9.37	8.81	14.50	6.49	13.40
7	KAJ Roach	12.0	77	75.18	5	6.24	9.27	9.11	15.40	6.42	14.40
8	A Nehra	20.0	156	157.23	10	13.29	9.26	9.30	15.60	7.80	12.00
9	M Morkel	15.0	119	123.55	8	10.02	9.33	9.56	14.88	7.93	11.25
10	SPD Smith	23.0	163	164.77	11	12.73	9.63	9.71	14.82	7.09	12.55
11	Saeed Ajmal	22.2	169	168.84	11	13.06	9.78	9.77	15.36	7.57	12.18
12	N McCullum	19.0	124	124.17	7	8.63	10.05	10.06	17.71	6.53	16.29
13	SCJ Broad	20.5	140	133.31	8	9.07	10.48	10.11	17.50	6.72	15.62
14	R Sidebottom	21.3	160	163.55	10	11.69	10.07	10.22	16.00	7.44	12.90
15	NO Miller	12.0	63	54.79	2	2.14	11.80	10.42	31.50	5.25	36.00
16	AD Mathews	12.0	83	82.46	4	5.35	10.59	10.54	20.75	6.92	18.00
17	M Aamer	23.0	152	142.74	8	8.54	11.14	10.61	19.00	6.61	17.25

O = No. of overs R = No. of runs R# = Adjusted no. of runs W = No. of wickets W* = Sum of weights of wickets
CBR* = Combined bowling rate CBR# = Combined bowling rate adjusted A = Bowling average E = Economy rate S = Strike rate

Continued on next page…

TABLE 2 (cont.)

Rank	Name	O	R	R#	W	W*	CBR*	CBR#	A	E	S
18	DL Vettori	19.1	109	101.60	3	2.94	13.13	12.34	36.33	5.69	38.34
19	DW Steyn	19.0	131	136.07	5	6.41	11.99	12.35	26.20	6.89	22.80
20	MH Yardy	20.0	136	133.73	4	5.36	12.98	12.80	34.00	6.80	30.00
21	JH Kallis	15.0	96	87.57	2	2.70	13.99	12.92	48.00	6.40	45.00
22	SJ Benn	12.0	78	76.15	2	2.73	13.23	12.97	39.00	6.50	36.00
23	SL Malinga	20.5	165	174.10	6	7.76	12.75	13.26	27.50	7.92	20.83
24	DJ Bravo	14.0	116	123.26	5	5.53	12.81	13.38	23.20	8.29	16.80
25	SE Bond	19.0	145	154.22	5	6.52	12.86	13.47	29.00	7.63	22.80
26	BAW Mendis	18.0	128	132.94	4	4.75	13.53	13.95	32.00	7.11	27.00
27	YK Pathan	16.0	140	146.14	4	5.43	14.31	14.76	35.00	8.75	24.00
28	Shahid Afridi	24.0	182	180.24	4	5.06	15.40	15.28	45.50	7.58	36.00
29	TT Bresnan	22.0	158	162.75	3	3.34	16.16	16.57	52.67	7.18	44.00
30	Harb. Singh	20.0	123	112.40	0	0	18.45	16.86	-	6.15	-
31	M Hafeez	14.0	123	133.32	2	2.70	17.87	19.06	61.50	8.79	42.00
32	RA Jadeja	12.0	117	141.34	2	2.85	18.02	20.74	58.50	9.75	36.00
33	SR Watson	16.0	163	188.54	2	0.78	27.01	30.89	81.50	10.19	48.00

O = No. of overs R = No. of runs R# = Adjusted no. of runs W = No. of wickets taken W* = Sum of weights of wickets taken
CBR* = Combined bowling rate CBR# = Combined bowling rate adjusted A = Bowling average E = Economy rate S = Strike rate

These and other cases clearly illustrate the importance of using the match conditions adjustment method. It not only gives a more accurate assessment of a bowler's performance, but can also give rise to a different ranking of the bowlers. Tait ranked 10th according to CBR*, but 5th according to CBR$^{\#}$. Kallis and Vettori also ranked 5 positions better after taking match conditions into account. On the other hand, CBR$^{\#}$ ranked Malinga, Bravo and Bond 4 positions lower than CBR* because they had generally bowled less economically per match than the group as a whole. There is obviously very high correlation between CBR$^{\#}$, CBR* and CBR. The ordinary correlation between CBR$^{\#}$ and CBR* is 0.98 and the Spearman correlation is 0.97. Despite these high correlations CBR$^{\#}$ remains the appropriate measure to use in a series of matches because the numerical values of CBR$^{\#}$ and CBR* can differ markedly.

DISCUSSION

At the end of a World Cup or other series a player is appointed as the man of the series. It is desirable that the panel should be able to determine which bowler (and batsman) was the best, because a medal, an amount of money and much prestige are involved. It can be very helpful to calculate the CBR$^{\#}$ values just before (or immediately after) the final match in order to identify the best bowler objectively.

The calculation of W*, the sum of the weights of the wickets taken by each bowler, and of R$^{\#}$, the sum of the adjusted number of runs, can easily be performed by working on the scorecards of the individual matches. It is then easy to calculate CBR$^{\#}$ for each bowler. The determination of wicket weights specifically for Twenty20 matches has to wait until a sufficient number of batsmen have played at least 20 international Twenty20 matches, the minimum requirement (Lemmer, 2005) for such a study. To date, a very small number of players meet this requirement. In order to be fair to players, special attention should be paid to the question of the most reliable performance measure in the case of a small number of scores.

The measure CBR$^{\#}$ has been designed specifically to determine which bowler has performed best in the series. It cannot be used to say that this bowler is currently the best. Short-term form is highly variable, and the same bowler can be far down in the ranking of the next series. For bowler rankings based on career data, the measures CBR of Lemmer (2002) and CBRW of Lemmer (2007) are much more appropriate. These measures are based on large data sets where influences of pitch and weather conditions tend to balance out.

CONCLUSION

The traditional and other bowling performance measures make no provision for the case of a small number of matches. If the bowling performances of bowlers have to be compared after a series (or a small number of matches), it is essential that the most suitable measure should be used. The use of wicket weights in Lemmer (2005) was the first step in this direction. The use of the individual match approach to address the problem of differing match conditions in the present study is the second step, and has led to the measure CBR$^{\#}$. Besides its theoretical justification of making the playing field more even, as far as match-specific conditions are concerned, its application to the bowling performances in the ICC Twenty20 Series of 2010, has shown that it makes a difference to the rankings.

REFERENCES

BAIRAM, E.I.; HOWELLS, J.M. & TURNER, G.M. (1990). Production functions in cricket: The Australian and New Zealand experience. *Applied Economics,* 22: 871-879.

BARR, G.D.I. & KANTOR, B.S. (2004). A criterion for comparing and selecting batsmen in limited overs cricket. *Journal of the Operational Research Society,* 55: 1266-1274.

BARR, G.D.I.; HOLDSWORTH, G.C. & KANTOR, B.S. (2008). Evaluating performances at the 2007 cricket world cup. *South African Statistical Journal,* 42(2): 125-142.

BASEVI, T. & BINOY, G. (2007). "The world's best Twenty20 players". [http://content-rsa.cricinfo.com/columns/content/story/311962.html]. Retrieved on 3 October 2007.

BEAUDOIN, D. & SWARTZ, T.B. (2003). The best batsmen and bowlers in one day cricket. *South African Statistical Journal,* 37(2): 203–222.

BRADMAN, D.G. (1950). *Farewell to cricket.* London: Hodder and Stoughton.

COHEN G.L. (2002). Cricketing chances. In G. Cohen & T. Langtry (Eds.), *Proceedings of the Sixth Australian Conference on Mathematics and Computers in Sport* (1-13). Sydney: University of Technology, NSW, Australia.

CRICINFO (2010a). "Cricket Scorecard". [http://www.cricinfo.com/world-twenty20-2010/engine/match/412703.html]. Retrieved on 15 June 2010.

CRICINFO (2010b). "ICC World Twenty20 2010". [http://www.cricinfo.com/world-twenty20/engine/records/averages/bowling.html]. Retrieved on 2 June 2010.

CROUCHER, J.S. (2000). Player ratings in one-day cricket. In G. Cohen & T. Langtry (Eds.), Proceedings of the Fifth Australian Conference on Mathematics and Computers in Sport (95-106). Sydney: University of Technology, NSW, Australia.

DUCKWORTH, F. & LEWIS, T. (2002). Review of the application of the Duckworth/Lewis method of target resetting in one-day cricket. In G. Cohen & T. Langtry (Eds.), *Proceedings of the Sixth Australian Conference on Mathematics and Computers in Sport* (127-140), Sydney: University of Technology, NSW, Australia.

KIMBER, A.C. (1993). A graphical display for comparing bowlers in cricket. *Teaching Statistics,* 15(3): 84–86.

LEMMER, H.H. (2002). The combined bowling rate as a measure of bowling performance in cricket. *South African Journal for Research in Sport, Physical Education and Recreation,* 24(2): 37-44.

LEMMER, H.H. (2004). A measure for the batting performance of cricket players. *South African Journal for Research in Sport, Physical Education and Recreation,* 26(1): 55-64.

LEMMER, H.H. (2005). A method for the comparison of the bowling performances of bowlers in a match or a series of matches. *South African Journal for Research in Sport, Physical Education and Recreation*, 27(1): 91-103.

LEMMER, H.H. (2007). The allocation of weights in the calculation of batting and bowling performance measures. *South African Journal for Research in Sport, Physical Education and Recreation,* 29(2): 75-85.

LEMMER, H.H. (2008a). An analysis of players' performances in the first cricket Twenty20 World Cup Series. *South African Journal for Research in Sport, Physical Education and Recreation,* 30(2): 73-79.

LEMMER, H.H. (2008b). Measures of batting performance in a short series of cricket matches. *South African Statistical Journal,* 42(1): 83-105.

O'RILEY, B.J. & OVENS, M. (2006). Impress your friends and predict the final score. *Journal of Sport Science and Medicine,* 5: 488-494.

SKINNER, I. (1995). The value of graphical presentation of data: An example from cricket. *Reflections on Sport and Mathematics,* 20: 1-6.

VAN STADEN, P.J. (2009). Comparison of crickets' bowling and batting performance using graphical displays. *Current Science,* 96(6): 764–766.

Prof. Hermanus H. LEMMER: 122 Fourth Avenue, Fairland 2170, Johannesburg, Republic of South Africa. Tel.: (w) +27 (0)11 678 3031, Tel.: (h), Fax.: +27 (0)11 678 4413, Cell.: +27 073 166 1934, E-mail: hoffiel@uj.ac.za

(Subject Editor: Dr. Richard Stretch)