

Comparison of cricketers' bowling and batting performances using graphical displays

Historically the principle criterion used for rating and comparing bowlers in the game of cricket has been the bowling average, calculated by dividing the number of runs conceded in a match (or a series of matches) by the number of wickets taken in the match(es):

$$AV = \frac{\text{Number of runs}}{\text{Number of wickets}} \quad (1)$$

Two additional performance criteria are also usually quoted. A bowler's economy rate is defined as the number of runs conceded per k balls bowled and is calculated by:

$$ER_k = k \times \frac{\text{Number of runs}}{\text{Number of balls}}, \quad (2)$$

where k is often chosen to be 6, so that ER_6 then denotes the runs per over (6 balls). Another popular choice for k is 100. The third criterion is the bowler's strike rate, originally proposed by Donald Bradman¹, which is given by the number of balls bowled divided by the number of wickets taken:

$$SR = \frac{\text{Number of balls}}{\text{Number of wickets}} \quad (3)$$

In the rest of this correspondence the three criteria will be referred to as the runs per wicket ratio, the runs per k balls ratio and the balls per wicket ratio, and denoted by:

$$RpW = \frac{\text{Number of runs}}{\text{Number of wickets}}, \quad (4)$$

$$RpB_k = k \times \frac{\text{Number of runs}}{\text{Number of balls}}, \quad (5)$$

$$BpW = \frac{\text{Number of balls}}{\text{Number of wickets}} \quad (6)$$

These three criteria can also be defined for the batsmen; so using standard terminology and notation will be beneficial.

A graphical representation for depicting all three bowling criteria has been proposed by Kimber², but does not seem to be widely used in the print or electronic media or in technical papers on cricket. This is surprising, since Kimber's graph is a simple yet powerful tool for comparing the performances of bowl-

ers. Therefore, the construction and interpretation of the graph for bowlers will be considered below. Use of the graph will then be extended by adapting it for a comparison of the batting and all-round performances of cricketers.

Any software with basic graphical capabilities should be suitable for constructing the graphs. In this correspondence the graphs were created with R, an open source environment and language for statistical computing and graphics³.

Although the graphs are applicable to any format of the game, bowling and batting records for players competing in the Indian Premier League (IPL) in 2008, obtained from the Cricinfo website⁴, have been utilized to illustrate the use and interpretation of the graphs. The IPL is played under the Twenty20 (or T20) format of cricket in which each team is given a single innings with a maximum of 20 overs. Twelve bowlers and 12 batsmen were selected – see Table 1 for their bowling and batting records. The selected bowlers all bowled at least 100 balls and took at least four wickets. Similarly, each selected batsman faced at least 100 balls and had at least four completed innings, where a completed innings is defined as an innings in which the batsman has been dismissed. Note that six cricketers were selected as bowlers and as batsmen.

From eqs (4)–(6) it follows that a hyperbolic relation exists between the three criteria:

$$RpB_k \times BpW = k \times RpW \quad (7)$$

Kimber suggested that the criteria can be represented graphically by plotting BpW against RpB_k on a scatter plot and augmenting the plot by adding hyperbolic contours representing RpW . Although Kimber used $k = 100$ for RpB_k , $k = 6$ or any other logical value of k can also be used.

Figure 1 shows the graph for the 12 selected bowlers from the IPL. A bowler would ideally like to simultaneously maximize the number of wickets taken and minimize the number of runs conceded, relative to the number of balls bowled. So it follows from eqs (4)–(6) that the better bowlers will tend to have lower values of RpW , RpB_k and BpW

and hence they should appear towards the lower left-hand corner of the graph.

Sohail Tanvir was the most prolific wicket taker in the IPL and therefore the IPL Purple Cap Winner (the IPL's version of the 'Bowler of the series' award). He had the lowest runs per wicket ratio and the second lowest runs per 100 balls and balls per wicket ratios. Amit Mishra was the bowler with the lowest balls per wicket ratio. Interestingly, the most economic bowler in the IPL was a cricketer more renowned for his batting ability, namely S. C. Ganguly. G. D. McGrath and S. M. Pollock were also economical, but they did not take a lot of wickets relative to the number of balls that they bowled. Hence their balls per wicket ratio and runs per wicket ratio were not that low. Their bowling records are typical of limited-overs bowlers, whose main task in the team's bowling squad is to restrict the run scoring of the opposing batsmen. Mohammad Asif and J. H. Kallis were two of the most expensive bowlers in the IPL, both conceding more than 150 runs per 100 balls (nine runs per over). Incidentally, they were also two of the more expensive players in the IPL, costing US\$ 650,000 and US\$ 900,000 respectively. Since Kallis only took four wickets, his balls per wicket and runs per wicket ratios were also extremely high.

Unfortunately, it may happen in the graph that bowlers with similar ratios are plotted over each other. In Figure 1, this happens with R. P. Singh and W. P. U. J. C. Vaas. Singh took three times more

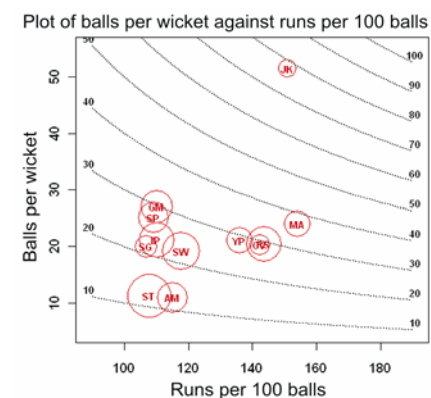


Figure 1. Comparison of the bowling performances of 12 bowlers in the IPL in 2008.

Table 1. Cricket records of 12 bowlers and 12 batsmen from the IPL in 2008

Player	Code	Team*	Country [†]	Balls	Runs	Wickets	RpB ₁₀₀	BpW	RpW
Bowlers									
S. C. Ganguly [‡]	SG	KKR	IND	120	128	6	106.67	20.00	21.33
J. H. Kallis [‡]	JK	BRC	SA	206	311	4	150.97	51.50	77.75
G. D. McGrath	GM	DD	AUS	324	357	12	110.19	27.00	29.75
Amit Mishra	AM	DD	IND	120	138	11	115.00	10.91	12.55
Mohammad Asif	MA	DD	PAK	192	296	8	154.17	24.00	37.00
I. K. Pathan [‡]	IP	KP	IND	318	350	15	110.06	21.20	23.33
Y. K. Pathan [‡]	YP	RR	IND	169	230	8	136.09	21.13	28.75
S. M. Pollock [‡]	SP	MI	SA	276	301	11	109.06	25.09	27.36
R. P. Singh	RS	DC	IND	308	442	15	143.51	20.53	29.47
Sohail Tanvir	ST	RR	PAK	247	266	22	107.69	11.23	12.09
W. P. U. J. C. Vaas	CV	DC	SL	102	145	5	142.16	20.40	29.00
S. R. Watson [‡]	SW	RR	AUS	325	383	17	117.85	19.12	22.53
Batsmen									
M. S. Dhoni	MD	CSK	IND	310	414	10	133.55	31.00	41.40
G. Gambhir	GG	DD	IND	379	534	13	140.90	29.15	41.08
S. C. Ganguly [‡]	SG	KKR	IND	307	349	12	113.68	25.58	29.08
A. C. Gilchrist	AG	DC	AUS	318	436	13	137.11	24.46	33.54
J. H. Kallis [‡]	JK	BRC	SA	183	199	11	108.74	16.64	18.09
S. E. Marsh	SM	KP	AUS	441	616	9	139.68	49.00	68.44
I. K. Pathan [‡]	IP	KP	IND	116	131	6	112.93	19.33	21.83
Y. K. Pathan [‡]	YP	RR	IND	243	435	14	179.01	17.36	31.07
S. M. Pollock [‡]	SP	MI	SA	111	147	8	132.43	13.88	18.38
V. Sehwag	VS	DD	IND	220	406	12	184.55	18.33	33.83
G. C. Smith	GS	RR	SA	362	441	9	121.82	40.22	49.00
S. R. Watson [‡]	SW	RR	AUS	311	472	10	151.77	31.10	47.20

*BRC, Bangalore Royal Challengers; CSK, Chennai Super Kings; DC, Deccan Chargers; DD, Delhi Daredevils; KKR, Kolkata Knight Riders; KP, Kings XI Punjab; MI, Mumbai Indians, and RR, Rajasthan Royals.

[†]AUS, Australia; IND, India; PAK, Pakistan; SA, South Africa, and SL, Sri Lanka.

[‡]Cricketers included as bowlers and as batsmen.

wickets than Vaas and he bowled approximately three times more balls and conceded approximately three times more runs than Vaas. This highlights the importance of taking the number of wickets into account when comparing bowlers. In Figure 1 this is accomplished by adding circles to the plot with radii relative to the number of wickets taken – this feature was not part of Kimber's originally proposed graph.

The number of balls bowled, the number of runs conceded and the number of wickets taken have traditionally always been part of the standard records kept and reported for bowlers, enabling the calculation of all three bowling criteria. For batsmen, however, until the early 1990s, only the total number of innings, the number of not-out innings (innings in which the batsmen were not dismissed) and the number of runs scored were reported. Due to this limited information, the batting average used to be the only batting criterion available. The batting average is defined as the number of runs scored in all innings divided by the number of completed innings:

$$AV = \frac{\text{Number of runs}}{\text{Number of completed innings}}. \quad (8)$$

Since the number of completed innings of a batsman can be interpreted as the number of times the wicket of the batsman has been taken, it follows that the batting average is also given by the runs per wicket ratio in eq. (4). Thus, the batting average for batsmen is equivalent to the bowling average for bowlers and both these averages are simply the runs per wicket ratio.

From the beginning of the 1990s, the number of balls faced by batsmen has been included in their batting records, enabling the rate at which they accumulate runs to be measured. The strike rate of a batsman is defined as the number of runs scored per k balls and is calculated by:

$$SR_k = k \times \frac{\text{Number of runs}}{\text{Number of balls}}, \quad (9)$$

where k is usually taken to be 100. Unfortunately the strike rates of bowlers and batsmen are not equivalent criteria, mak-

ing the terminology somewhat ambiguous. Instead, comparing eq. (9) with eq. (2), we notice that the strike rate of batsmen is equivalent to the economy rate of bowlers. To avoid confusion, the term runs per k balls ratio and eq. (5) will be used for batsmen as was done for the bowlers before.

Currently, the runs per wicket and the runs per k balls ratios are the only two performance criteria commonly used for comparing the batting abilities of cricketers. In order to construct a graph for batsmen analogous to that for bowlers, a third criterion is needed. Fortunately a third criterion is hiding in the data. Recall that the third criterion for bowlers is the balls per wicket ratio, given in eq. (6). For a batsman this criterion can also be calculated, since the number of balls faced is available, as is the number of times the batsman's wicket has been taken. For everyday referral in general cricket terminology, it is suggested that this new criterion for batsmen be called the survival rate, since it can be viewed as a measure of the ability of a batsman to survive the opposition's bowling

attack and defend his wicket. However, for uniformity in this correspondence, the third criterion will be referred to as the balls per wicket ratio as was done for bowlers.

Given the three criteria, construction of the graph for the batsmen proceeds in exactly the same way as for the bowlers: BpW is plotted against RpB_k on a scatter plot and the plot was then augmented by adding hyperbolic contours representing RpW . To take account of the number of times each batsman was dismissed, circles can be added to the plot with radii relative to the number of times each batsman's wicket has been taken. There is of course one important difference between the three criteria for bowlers and batsmen. Whereas for bowlers small values for RpW , RpB_k and BpW are preferable, batsmen would like to maximize these values by scoring as many runs as possible and losing their wickets as seldom as possible, relative to the

number of balls faced. Thus better batsmen will tend to appear towards the upper right-hand corner of the graph.

In Figure 2 the various ratios of the 12 selected batsmen in Table 1 are represented. S. E. Marsh was the IPL Orange Cap Winner (the IPL's version of the 'Batsman of the series' award) for scoring the most runs. He also had by far the highest runs per wicket and balls per wicket ratios among all the batsmen. At a cost of just US\$ 30,000, Marsh was considered by many cricket analysts as the best value-for-money player in the IPL. G. C. Smith was another batsman with high runs per wicket and balls per wicket ratios. However, his runs per 100 balls ratio was rather low in terms of T20 cricket. A high balls per wicket ratio combined with a low runs per 100 balls ratio is typical of a relatively more defensive batsman. Conversely, relatively more offensive batsmen, like for example Y. K. Pathan and V. Sehwag, will have high runs per 100 balls ratio and low balls per wicket ratio. Most batsmen of course fall between these two extremes, in that they manage to protect their wickets while still accumulating runs at a reasonably fast rate. Examples from the IPL are M. S. Dhoni, G. Gambhir, A. C. Gilchrist and S. R. Watson.

The all-round performance of those cricketers who bat and bowl regularly can be analysed by plotting their three bowling criteria and their corresponding three batting criteria on the same scatter plot. This has been done in Figure 3 for the six cricketers in Table 1, whose bowling and batting records are provided. Watson, who was named 'Player of the series' in the IPL, had the ideal all-round performance in that his three batting criteria all had higher values than his corresponding three bowling criteria. Although regarded in the cricket fraternity as an excellent all-rounder, Kallis did not perform well in the IPL as batsman or as bowler. His three batting criteria all had much lower values than the corresponding bowling criteria.

The rate at which Y. K. Pathan scored his runs was higher than the rate at which he conceded runs as a bowler. Also, his runs per wicket ratio was higher for batting than for bowling. Only with respect to his balls per wicket ratio was Pathan's value for bowling higher than his value for batting. His all-round performance in

the IPL was typical of a batting all-rounder who bowls occasionally. Similar to Y. K. Pathan, Pollock scored runs at a faster rate than the rate at which he conceded runs, while his balls per wicket ratio was higher for bowling than for batting. However, contrary to Y. K. Pathan, Pollock's runs per wicket ratio was higher for bowling than for batting. Pollock's all-round performance in the IPL was typical of a bowling all-rounder who bats lower down the order, that is, not in the top six of the batting line-up. The three batting ratios for I. K. Pathan were approximately the same as his three bowling ratios. He bowled 318 balls in the IPL and faced 116 balls as batsman, so he was mainly used as a bowler by his team. If it though happened that the number of balls he bowled was approximately equal to the number of balls he faced as a batsman, then his batting performance would have cancelled out his bowling performance (or vice versa). It is debatable whether a cricketer like this is beneficial to the team.

A simple way of graphically comparing the bowling and batting performances of cricketers was illustrated using records from the IPL. The graphs are applicable to any format of cricket and can furthermore be used to identify different types of players, for example, offensive batsmen, bowling all-rounders, etc. The use of the graphs can be extended in numerous ways – see van Staden⁵ for some interesting examples.

Plot of balls per wicket against runs per 100 balls

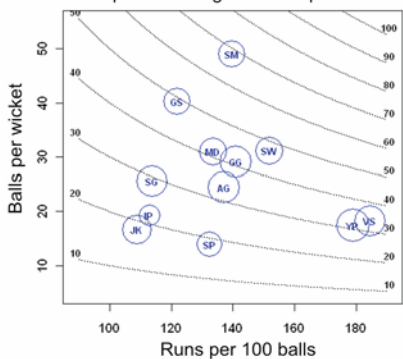


Figure 2. Comparison of the batting performances of 12 batsmen in the IPL in 2008.

Plot of balls per wicket against runs per 100 balls

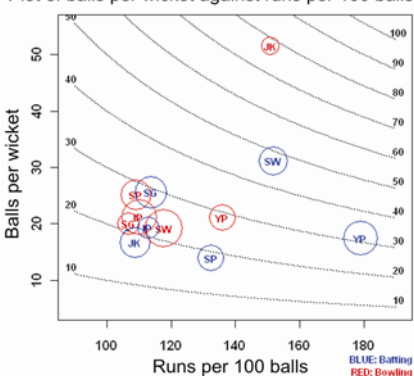


Figure 3. Comparison of the batting and bowling performances of six cricketers in the IPL in 2008.

1. Bradman, D. G., *Farewell to Cricket*, Hodder and Stoughton, London, 1950, p. 118.
2. Kimber, A., *Teach. Stat.*, 1993, **15**, 84–86.
3. The Comprehensive R Archive Network, <http://cran.r-project.org/>
4. Cricinfo: The home of cricket, <http://www.cricinfo.com/>
5. van Staden, P. J., Technical Report 08/01, Department of Statistics, University of Pretoria, South Africa, 2008.

Received 19 June 2008; revised accepted 11 February 2009

PAUL J. VAN STADEN

Department of Statistics,
University of Pretoria,
Pretoria, South Africa 0002
e-mail: paul.vanstaden@up.ac.za