UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

FACULTY OF NATURAL AND AGRICULTURAL SCIENCES

DEPARTMENT OF STATISTICS

Comparison of bowlers, batsmen and all-rounders in cricket using graphical displays

Paul J. van Staden

TECHNICAL REPORT
08/01

Comparison of bowlers, batsmen and all-rounders in cricket using graphical displays

Paul J. van Staden

2008

UNIVERSITY OF PRETORIA

FACULTY OF NATURAL AND AGRICULTURAL SCIENCES

DEPARTMENT OF STATISTICS

ISBN: 978-1-86854-733-3

Comparison of bowlers, batsmen and all-rounders in cricket using graphical displays

Paul J. van Staden

Department of Statistics, University of Pretoria, Pretoria, 0002, South Africa

e-mail: *paul.vanstaden@up.ac.za*

Summary: *In cricket the comparison of cricketers' batting and bowling abilities is usually done with very basic performance measures. More sophisticated measures have been proposed, but are generally not used due to a variety of reasons, including the statistical illiteracy of those involved in cricket, the way cricket data is captured and presented for bowlers and for batsmen and the different rules applicable for the various formats of the game. Graphical displays for comparisons have not featured prominently. In this paper a graph, originally proposed for comparing bowlers, is presented and adapted for comparing batsmen and all-rounders. The construction and interpretation of the graphs are illustrated with cricket records from the recent Indian Premier League (IPL).*

Keywords: Cricket; batting, bowling and all-round performance criteria; graphical displays.

HISTORICALLY the principle criterion used for rating and comparing bowlers in the game of cricket has been the so-called bowling average. The average for a bowler is 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),

$$AV = \frac{\text{Number of Runs}}{\text{Number of Wickets}}. \qquad (1)$$

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

$$ER_k = k \times \frac{\text{Number of Runs}}{\text{Number of Balls}}, \qquad (2)$$

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 Sir Donald Bradman[1], which is given by the number of balls bowled divided by the number of wickets taken,

$$SR = \frac{\text{Number of Balls}}{\text{Number of Wickets}}. \tag{3}$$

Although all three criteria are usually quoted in the cricket records of bowlers, the strike rate is seldom interpreted and bowlers are usually rated solely on their bowling average and economy rate.

In cricket terminology the three criteria are usually referred to as the bowling average, the economy rate and the strike rate and denoted by AV, ER_k and SR respectively. However, in the rest of this paper 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

$$RpW = \frac{\text{Number of Runs}}{\text{Number of Wickets}}, \tag{4}$$

$$RpB_k = k \times \frac{\text{Number of Runs}}{\text{Number of Balls}}, \tag{5}$$

$$BpW = \frac{\text{Number of Balls}}{\text{Number of Wickets}}. \tag{6}$$

The same three criteria can also be defined for batsmen, so using standard terminology and notation will be beneficial. This will furthermore allow for the comparison of cricketers' bowling and batting abilities and, by doing so, the identification of the best all-rounders (players that excel in bowling and batting).

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. Hence it follows from equations (4) to (6) that the better bowlers in cricket will tend to have lower values of RpW, RpB_k and BpW.

The three criteria are not infallible. It is immediately clear that if a bowler does not take any wickets, then only RpB_k can be calculated. Also, if the number of wickets taken by a bowler is small in magnitude relative to the number of balls bowled or the number of runs conceded, then any increase in the number of wickets taken may lead to large decreases in the values of BpW and RpW, even though the bowling ability of the bowler has not improved significantly. To illustrate, suppose a bowler has bowled a total of 80 balls, conceded 60 runs and has taken only 2 wickets

so that $BpW = \frac{80}{2} = 40$ and $RpW = \frac{60}{2} = 30$. If the bowler takes a wicket with the next ball bowled (no runs obviously conceded), then $BpW = \frac{81}{3} = 27$ and $RpW = \frac{60}{3} = 20$. Therefore the number of wickets itself – in effect, the denominator in the balls per wicket and the runs per wicket ratios – should also be considered when evaluating a bowler's performance. In fact, as will be shown in the next example, using only ratios to compare bowlers can lead to spurious outcomes.

Consider a simple example with hypothetical data that conjures up some startling results. Suppose two bowlers competed for the same team in the Indian Premier League (IPL) and that their team reached and played in the final. Their bowling records are given in Table 1. The records are divided into three stages: firstly all the team's matches up to but excluding the final of the IPL, secondly the final itself and thirdly all the team's matches in the IPL including the final. The lowest values of RpB_{100}, BpW and RpW between the two bowlers at each stage are underlined. The performances of the two bowlers were very similar in the matches up to the final, with Bowler B slightly more economic (he had a lower value for RpB_{100}), and Bowler A having had slightly lower values for BpW and RpW. In the final itself Bowler B was again more economic, while Bowler A once again had the lower values for BpW and RpW. We would therefore suspect that Bowler B was the most economic of these two bowlers in the IPL, while Bowler A had the lower balls per wicket and runs per wicket ratios. However, looking at the last two rows of Table 1, we notice that the order of each criterion is reversed! In effect, Bowler A actually had the lowest value for RpB_{100}, while Bowler B had the lowest values for BpW and RpW. Although mathematical trickery might be suspected, there is in fact a straightforward mathematical explanation[2]. When comparing ratios, then

$$\frac{a_1}{b_1} < \frac{c_1}{d_1} \text{ and } \frac{a_2}{b_2} < \frac{c_2}{d_2} \not\Rightarrow \frac{a_1 + a_2}{b_1 + b_2} < \frac{c_1 + c_2}{d_1 + d_2}. \qquad (7)$$

Bolt[3] explained this concept using simple geometric illustrations applied to runs per wicket ratios and concluded that:

> "This method of illustration draws attention to the fact that bowling averages are really rates and that, just as when considering average speeds we need to take account of the time for which each speed is maintained, so when considering bowling averages we need to take account of the number of wickets for which each average has been calculated."

His conclusion is naturally also true for the runs per k balls and the balls per wicket ratios. The above example therefore demonstrates that it is not wise to compare bowlers' performances using only ratios, as is often done by cricket statisticians, analysts, commentators and other role-players in the cricket fraternity. The number of wickets taken and the number of balls bowled – in effect, the denominators of the various ratios – should also be taken into consideration.

Table 1. Hypothetical bowling records for two bowlers in the Indian Premier League (IPL)

Bowler	Stage of IPL*	Balls	Runs	Wickets	RpB_{100}	BpW	RpW
A	Before final	228	240	15	105.26	15.20	16.00
B	Before final	240	244	15	101.67	16.00	16.27
A	In final	12	24	3	200.00	4.00	8.00
B	In final	24	47	5	195.83	4.80	9.40
A	After final	240	264	18	110.00	13.33	14.67
B	After final	264	291	20	110.23	13.20	14.55

* Indicates which matches are included in the records:
Before final: All the team's matches up to but excluding the final.
In final: The final itself.
After final: All the team's matches including the final.

There have been attempts to combine the three criteria for bowlers into a single measure[4–6]. A graphical representation for depicting all three criteria has also been proposed by Kimber[7], 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 bowlers' performances. Therefore the construction and interpretation of the graph for bowlers will be considered below. The use of the graph will then be extended by adapting it for comparison of cricketers' batting abilities. Finally it will be illustrated how the graph can be used to compare all-rounders.

Any software with basic graphical capabilities should be suitable for constructing the graphs. In this paper the graphs were created with R, an open source environment and language for statistical computing and graphics[8].

The Indian Premier League (IPL) and Twenty20 (T20) cricket

Throughout the paper bowling and batting records for players competing in the IPL in 2008 will be used to illustrate the concepts and graphs. These records were obtained from the Cricinfo website[9]. The IPL is played under the so-called Twenty20 (or T20) format of cricket. In T20 cricket each team is given a single innings with a maximum

of 20 overs. Whereas a typical one-day game, with a maximum of 50 overs per innings, is scheduled for six hours, and a match consisting of at most two innings per team is usually scheduled for at least three days (for instance, a Test match between two countries is scheduled for five days), a T20 game is completed in about three hours. Thus, in terms of time duration, T20 cricket is closer to other popular team sports like football, hockey and rugby.

The reason for considering T20 cricket and the IPL in particular in this paper is twofold. Firstly, in most of the research articles published on player performance and the measuring thereof, researchers have focused on Test cricket and one-day cricket. This is not unexpected, since the inaugural Test matches were played between Australia and England in March and April of 1877 in Australia, while the first one-day international (ODI), also between Australia and England, took place on 5 January 1971 in Melbourne, Australia. Thus, there is a wealth of cricket data available for these two formats of cricket. T20 cricket was only launched in 2003 as a domestic inter-county competition by the England and Wales Cricket Board. The first ever T20 international, between New Zealand and Australia, was only played on 17 February 2005 in Auckland. It is therefore no coincidence that virtually no analysis has been done on the performances of cricketers in this format. The second reason for using the IPL in this paper is that, although the IPL is a domestic cricket league in India, the eight teams taking part in the league had, apart from the local Indian players, also the best available foreign players in their squads. The IPL therefore provides a wonderful opportunity for comparing cricket players from different countries with various skills. In competitions of the International Cricket Council (ICC), like the ICC Cricket World Cup and the ICC World Twenty20, the best countries and players of the world also compete, but in these competitions the number of games played by players is not enough to enable credible comparisons to be made.

A total of 163 cricketers played in the IPL, of whom 98 bowled at least one ball in the tournament. Recall that the values of BpW and RpW are adversely affected by a small number of wickets. Therefore only bowlers who took at least five wickets were considered for inclusion in the study, the exception being ST Jayasuriya and JH Kallis. They only took four wickets each, but are included, since they are also included as batsmen. A second requirement for inclusion was that bowlers had to have bowled a minimum of 100 balls in the tournament. Apart from Jayasuriya and Kallis,

44 bowlers complied with both these two requirements. It was decided to select 16 bowlers from the 46 candidates. These 16 bowlers along with their bowling records are listed in Table 2.

Table 2. Cricket records of 16 bowlers and 16 batsmen (including 8 all-rounders) from the IPL in 2008

Player	Code	Team*	Country†	Balls	Runs	Wickets	RpB_{100}	BpW	RpW
Bowlers:									
SC Ganguly‡	SG	KKR	IND	120	128	6	106.67	20.00	21.33
ST Jayasuriya‡	SJ	MI	SL	126	159	4	126.19	31.50	39.75
JH Kallis‡	JK	BRC	SA	206	311	4	150.97	51.50	77.75
GD McGrath	GM	DD	AUS	324	357	12	110.19	27.00	29.75
A 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
JA Morkel‡	JM	CSK	SA	288	399	17	138.54	16.94	23.47
IK Pathan‡	IP	KP	IND	318	350	15	110.06	21.20	23.33
YK Pathan‡	YP	RR	IND	169	230	8	136.09	21.13	28.75
SM Pollock‡	SP	MI	SA	276	301	11	109.06	25.09	27.36
RP 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
S Sreesanth	SS	KP	IND	307	442	19	143.97	16.16	23.26
Umar Gul	UG	KKR	PAK	135	184	12	136.30	11.25	15.33
WPUJC Vaas	CV	DC	SL	102	145	5	142.16	20.40	29.00
SR Watson‡	SW	RR	AUS	325	383	17	117.85	19.12	22.53
Batsmen:									
MS 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
SC Ganguly‡	SG	KKR	IND	307	349	12	113.68	25.58	29.08
AC Gilchrist	AG	DC	AUS	318	436	13	137.11	24.46	33.54
ST Jayasuriya‡	SJ	MI	SL	309	514	12	166.34	25.75	42.83
JH Kallis‡	JK	BRC	SA	183	199	11	108.74	16.64	18.09
BB McCullum	BM	KKR	NZ	92	188	3	204.35	30.67	62.67
SE Marsh	SM	KP	AUS	441	616	9	139.68	49.00	68.44
JA Morkel‡	JM	CSK	SA	163	241	7	147.85	23.29	34.43
IK Pathan‡	IP	KP	IND	116	131	6	112.93	19.33	21.83
YK Pathan‡	YP	RR	IND	243	435	14	179.01	17.36	31.07
SM 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
GC Smith	GS	RR	SA	362	441	9	121.82	40.22	49.00
SB Styris	SS	DC	NZ	123	112	6	91.06	20.50	18.67
SR Watson‡	SW	RR	AUS	311	472	10	151.77	31.10	47.20

* Teams:

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 — RR: Rajasthan Royals

† Countries:

AUS: Australia — IND: India

NZL: New Zealand — PAK: Pakistan

SA: South Africa — SL: Sri Lanka

‡ Indicates all-rounders.

With respect to batting, 154 cricketers faced at least one ball in the tournament. As with the bowlers, there were two requirements for inclusion in the study. Firstly, a batsman must have had at least 5 completed innings, where a completed innings is defined as an innings in which the batsman has been dismissed. Secondly, a batsman must have faced a minimum of 100 balls in the tournament. Only 46 batsmen complied with both requirements. As with the bowlers, 16 batsmen were selected – see Table 2 for these batsmen and their batting records. Although BB McCullum only played and batted in four matches, had only three completed innings and faced a total of just 92 balls, it was decided to include him as one of the 16 batsmen. The reason is that he scored the highest ever score in T20 history, 158 runs not out from just 73 balls, in the opening match of the IPL. Finally, note that eight cricketers, including Jayasuriya and Kallis, were selected as bowlers and as batsmen in order to be compared as all-rounders.

A graph for comparing bowlers

From equations (4) to (6) it follows that a hyperbolic relation exists between the three criteria,

$$RpB_k \times BpW = k \times RpW \,. \qquad (8)$$

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. Note that although Kimber used $k = 100$ for RpB_k, $k = 6$ or any other logical value of *k* can also be used. To explain the interpretation of the graph, consider three Indian bowlers: IK Pathan, RP Singh and S Sreesanth. Their ratios are represented in Figure 1 by their respective codes, "IP", "RS" and "SS". Recall that better bowlers tend to have lower values for RpW, RpB_k and BpW and hence they should appear towards the lower left-hand corner of the graph, as indicated by the arrows and the code "BEST" in Figure 1. Singh bowled just one more ball than Sreesanth, while they conceded the same amount of runs. Therefore they had near-identical runs per 100 balls ratios – see the horizontal axis of Figure 1. IK Pathan bowled just a few more balls than Singh and Sreesanth, but conceded nearly 100 runs less than them. Hence he was the most economic bowler of the three. IK Pathan and Singh took 15 wickets each. Because IK Pathan bowled a few balls more than Singh,

his balls per wicket ratio was slightly higher than Singh's – see the vertical axis of Figure 1. Sreesanth took four more wickets (19) than the other two bowlers, so his balls per wicket ratio was by far the lowest of the three bowlers. IK Pathan and Sreesanth had very similar runs per wicket ratios – see the hyperbolic contours indicated by the dotted lines – while Singh's runs per wicket ratio was higher. Clearly out of these three bowlers, Singh's bowling performance was the worse. Deciding whom of IK Pathan or Sreesanth was the better bowler, is not so straightforward. The decision depends on whether the capturing of wickets or the restriction of runs is the most important consideration – in limited-overs cricket the restriction of runs is usually the primary objective.

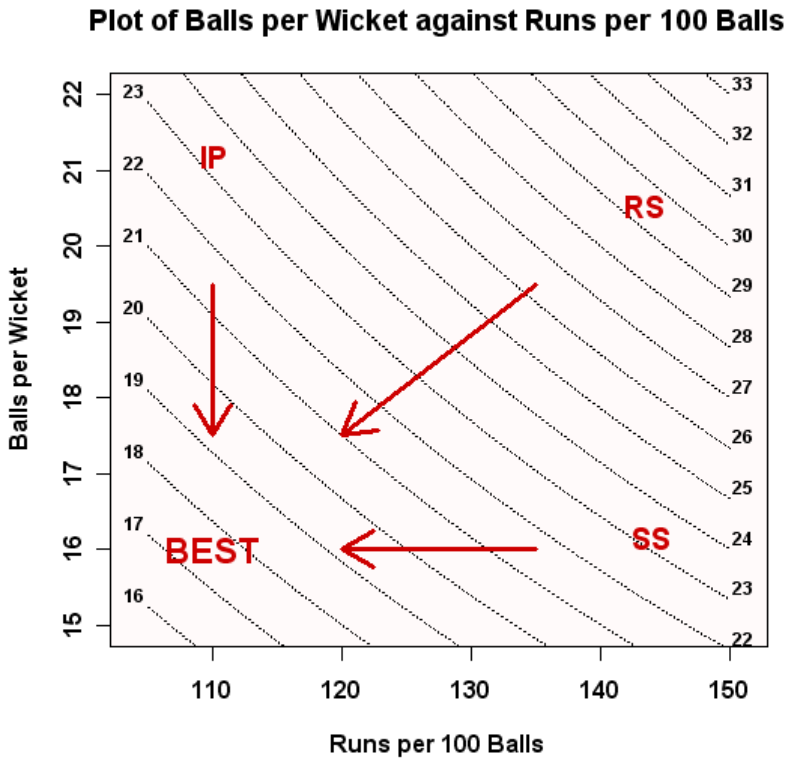

Figure 1. Comparison of the bowling performances of IK Pathan, RP Singh and S Sreesanth in the IPL in 2008

Figure 2 shows the graph for all 16 selected bowlers. Analogous to Kimber, $k = 100$ is again used. Sohail Tanvir was the most prolific wicket taker in the IPL with 22 victims and therefore was the IPL Purple Cap Winner (the IPL's version of the Bowler of the Series award). Of the 46 bowlers considered for inclusion in the study, Sohail Tanvir had the lowest runs per wicket ratio and the second lowest runs per 100 balls and balls per wicket ratios. A Mishra was the bowler with the lowest balls per wicket ratio. He, Sohail Tanvir and Umar Gul all had balls per wicket ratios

smaller than 12 balls per wicket. A bowler can bowl a maximum of four overs in a T20 match, implying a maximum of 24 deliveries. So it follows from their balls per wicket ratios that, given that they bowled their full complement of overs in each match, Mishra, Sohail Tanvir and Umar Gul would take on average more than two wickets per match.

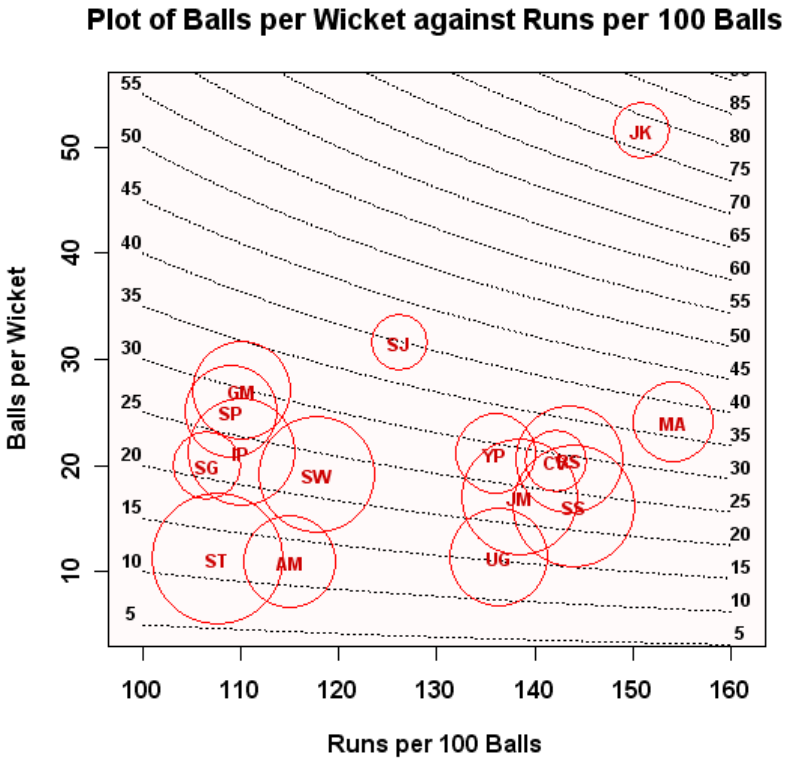

Figure 2. Comparison of the bowling performances of 16 bowlers in the IPL in 2008

Interestingly, the most economic bowler in the IPL was a cricketer more renowned for his batting ability, namely SC Ganguly. GD McGrath and SM Pollock were also very economic, but they did not take a lot of wickets relative to the number of balls they bowled. Hence their balls per wicket ratios and also their runs per wicket ratios 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 Kallis were two of the most expensive bowlers in the IPL, both conceding more than 150 runs per 100 balls (9 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 2 this happens with Singh and WPUJC Vaas. Careful examination of their bowling records in Table 2 reveals that Singh took three times more wickets than Vaas and that he bowled approximately three times more balls and conceded approximately three times more runs than Vaas. This highlights again the importance of taking the number of wickets into account when comparing bowlers. In Figure 2 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.

A graph for comparing batsmen

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 a batsman however, until the early 1990s, only the total number of innings, the number of so-called not out innings (innings in which the batsman was 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,

$$AV = \frac{\text{Number of Runs}}{\text{Number of Completed Innings}}. \qquad (9)$$

Often a batsman has a high batting average only because of many not out innings relative to the total number of innings. Therefore alternative estimates for the batting average have been proposed, for example estimates based upon the product limit estimator[10,11] and the first moment (in effect, the mean) of the gamma and the Weibull distributions[12,13]. However, in this study the batting average as given in equation (9) will be used. Since the number of completed innings of a batsman can be interpreted as the number of times that the wicket of the batsman has been taken, it follows that the batting average is also given by the runs per wicket ratio in equation (4). Thus, the batting average for batsmen is equivalent to the bowling average of bowlers and both these averages are simply the runs per wicket ratio.

From the beginning of the 1990s the number of balls faced by a batsman has been included in the batting records of batsmen. The reason for the inclusion is so that

the rate at which a batsman accumulates runs can also be measured. The so-called strike rate of a batsman is then defined as the number of runs scored per k balls and is calculated by

$$SR_k = k \times \frac{\text{Number of Runs}}{\text{Number of Balls}}, \qquad (10)$$

where k is usually taken to be 100. Unfortunately the strike rates of bowlers and batsmen are not equivalent criteria, making the terminology somewhat ambiguous. Instead, comparing equation (10) with equation (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 equation (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 cricketers' batting abilities. Some additional measures have been defined, for example the geometric coefficient[14], the coefficient of variation[15] and the consistency coefficient[16], which are all measures of the consistency of batsmen (consistency is related to the variability of batting scores). These measures of consistency have also been combined with the two commonly used criteria mentioned above to create new measures, for example the consistency-adjusted average[14] and also a single measure based upon the exponentially weighted batting average, the runs per 100 balls ratio and the consistency coefficient[16]. Another interesting method for comparing the batting ability of batsmen is the use of stochastic dominance rules[17].

In order to construct a graph for batsmen analogous to the graph for bowlers, a third criterion is needed. Fortunately a simple third criterion is hiding in the data. Recall that the third criterion for bowlers is the balls per wicket ratio, given in equation (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 is 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 paper, 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 batsmen proceeds exactly the same as for bowlers: BpW is plotted against RpB_k on a scatter plot and the plot is 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 for 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 3 the various ratios of the 16 selected batsmen in Table 2 are represented. SE Marsh was the IPL Orange Cap Winner (the IPL's version of the Batsman of the Series award) for scoring the most runs in the IPL (616 runs). Marsh also had by far the highest runs per wicket and balls per wicket ratios of 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. GC Smith was another batsman with very high runs per wicket and balls per wicket ratios. Note however that 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 YK Pathan and V Sehwag, will have very high runs per 100 balls ratios and low balls per wicket ratios. 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 MS Dhoni, G Gambhir and AC Gilchrist. Consider the three batsmen in Figure 3 with the lowest runs per wicket ratios: SB Styris, Kallis and Pollock. These three batsmen had similar runs per wicket ratios, but clearly Pollock was the most attacking batsman of the three. Styris was the only batsman who faced more than 100 balls in the IPL and scored runs at a rate of less than 100 runs per 100 balls. Another New Zealander, McCullum, managed to score runs at the unbelievable rate of more than 200 runs per 100 balls. This was mainly due to his first innings of 158 runs not out from 73 balls. In none of his other three innings McCullum could score more than 24 runs. This shows again that the number of wickets, which represents the

number of completed innings for batsmen, should be taken into account when comparing batsmen, since a single outstanding (or alternatively bad) performance can give a false picture of a batsman's ability.

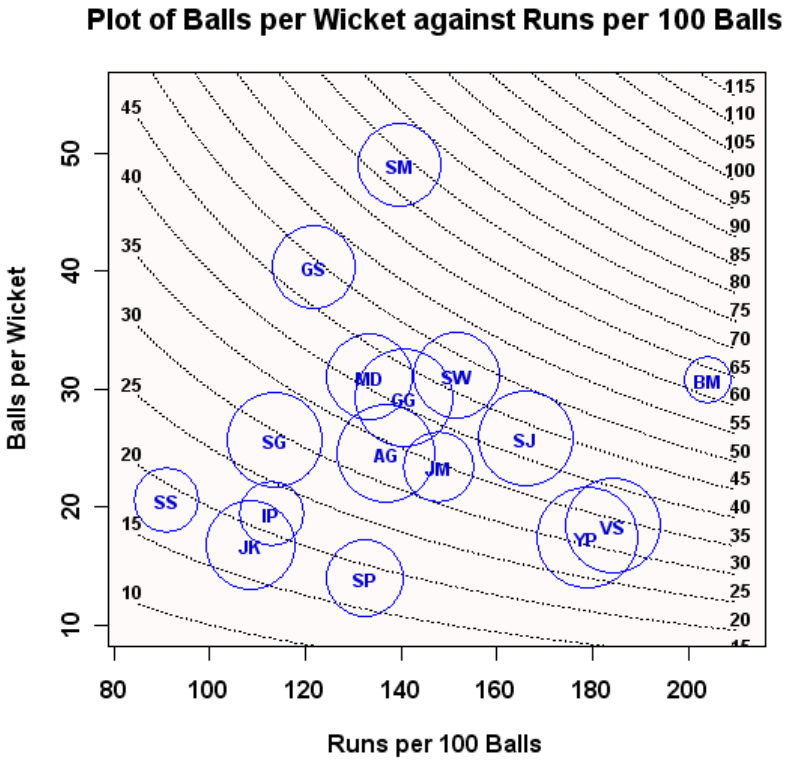

Figure 3. Comparison of the batting performances of 16 batsmen in the IPL in 2008

A graph for comparing all-rounders

Most bowlers must bat at some stage, whereas most batsmen only bowl occasionally. However, not all cricketers are specialists in batting and in bowling. Historically a cricket player was labeled an all-rounder if the player's batting and bowling abilities were of such a high level that a team would select the player solely as a batsman or as a bowler. Nowadays, in the professional era of cricket, true all-rounders are rare and any cricketer with decent batting and bowling records in Test cricket and/or limited-overs cricket is referred to as an all-rounder. A rule of thumb that is often used to identify a cricketer as an all-rounder, is that the runs per wicket ratio for batting should be higher than the corresponding ratio for bowling (in effect, the batting average should be higher than the bowling average). This does however not take the number of balls faced or the number of balls bowled into account.

Before proposing a graph for comparing all-rounders, other measures for comparing all-rounders will briefly be considered. Lewis[18] used the Duckworth-Lewis methodology[19,20] (the system currently utilized for adjusting cricket scores in interrupted matches in all limited-overs matches sanctioned by the ICC) to create new measures for batting and bowling performances in one-day cricket. He proposed the so-called net batting contribution, the net bowling contribution and, by adding these two measures, the total net contribution, which can be used to rank cricketers in a one-day match or a series of one-day matches. Since better batting and bowling performances are represented by larger values of the net batting and net bowling contributions, it follows that a cricketer who makes high positive contributions with respect to batting and bowling can be considered an all-rounder. For example, Lewis commented with respect to the Victoria Bitter (VB) Series in Australia in 2002/03, which he used to illustrate his measures, that the Australian cricketer DS Lehmann could be regarded as the best all-rounder in the tournament, since he had the highest total net contribution of all the players who had positive net contributions in both batting and bowling,. However, Lewis did not define a measure for the all-round contribution. Let Bat_C, $Bowl_C$ and $Total_C$ denote the net batting, net bowling and total net contribution for cricketers in one-day cricket as defined by Lewis. Then for cricketers with $Bat_C > 0$ and $Bowl_C > 0$, the all-round performance can be measured by the harmonic mean of Bat_C and $Bowl_C$, given by

$$AR_C = \frac{2 \times Bat_C \times Bowl_C}{Bat_C + Bowl_C} = \frac{2 \times Bat_C \times Bowl_C}{Total_C}, \qquad (11)$$

where AR_C can be interpreted as the all-round contribution. Note that, whereas $Total_C$ measures the contributions made by the batting *or* the bowling *or* the batting and the bowling of a cricketer, AR_C is indicative of the batting *and* the bowling contributions of the cricketer. For example, Lehmann had the highest value for AR_C in the VB Series, confirming the comment by Lewis. However, the highest value for $Total_C$ was obtained by his countryman, ML Hayden, which was the same value as for Bat_C, since Hayden did not bowl in the VB Series.

Beaudoin and Swartz[21] also used the Duckworth-Lewis methodology to develop a measure for evaluating the performances of batsmen and bowlers in one-day cricket. Their runs per match statistic is given by

$$RM = 100 \times \frac{\text{Number of Runs}}{\text{Resources Used}}, \qquad (12)$$

where the resources are defined according to the Duckworth-Lewis methodology. Good batsmen will have high RM batting statistics, while good bowlers will have low RM bowling statistics. Hence, as proposed by Beaudoin and Swartz, the performance of an all-rounder can be assessed by subtracting the all-rounder's RM bowling statistic from the RM batting statistic.

Johnston *et al.*[22] used a dynamic programming formulation, developed by Clarke[23], to obtain batting and bowling performance measures in one-day cricket. The sum of these two measures is then a measure of the total performance. Analogous to the measure proposed in equation (11), the harmonic mean of the batting and bowling performance measures can be interpreted as a measure of the all-round performance. Unfortunately the batting and bowling measures proposed by Lewis, by Beaudoin and Swartz and by Johnston *et al*. can only be used in limited-overs cricket and not in Test cricket, limiting the use of the proposed all-round performance measures.

A simple way of comparing the all-round performance of cricketers is to plot the three bowling criteria and the corresponding three batting criteria on the same scatter plot. This is done in Figure 4 for the 8 selected all-rounders from the IPL. Ganguly, JA Morkel and SR Watson had the ideal all-round performances in the IPL in that their three batting criteria all had higher values than their corresponding three bowling criteria. However, it is clear from Figure 4 that, while Ganguly performed brilliantly as bowler, his batting was not that outstanding. In contrast, Morkel and Watson excelled in terms of both batting and bowling, and Watson was named Player of the Series for his superb all-round performances. Although usually 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 Jayasuriya and YK Pathan scored their runs was higher than the rate at which they conceded runs. Also, their runs per wicket ratios for batting were higher than their runs per wicket ratios for bowling. Only with respect to their balls per wicket ratios were their values for bowling higher than their values for batting. Their all-round performances in the IPL were typical of so-called batting all-

rounders who bowl occasionally. Similar to Jayasuriya and YK Pathan, Pollock scored runs at a faster rate than the rate at which he conceded runs, while his balls per wicket ratio for bowling was higher than his balls per wicket ratio for batting. However, contrary to Jayasuriya and YK 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 IK Pathan were approximately the same as his three bowling ratios. From Table 2 it can be seen that he bowled 318 balls 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 batsman, then his batting performance would have cancelled out his bowling performance (or vice versa). It is debatable whether an all-rounder like this is then beneficial to the team.

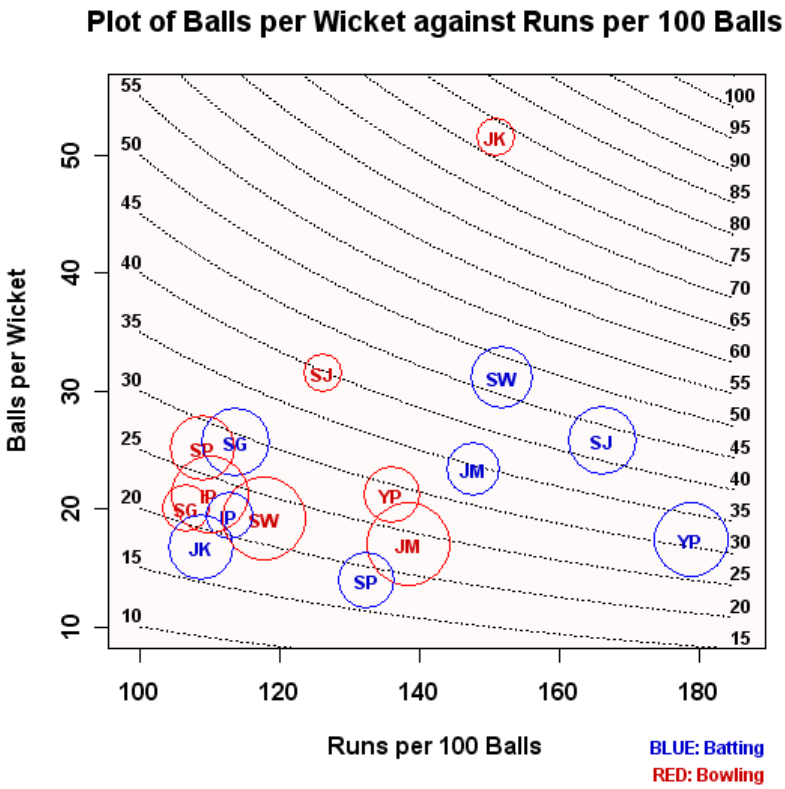

Figure 4. Comparison of the batting and bowling performances of 8 all-rounders in the IPL in 2008

Conclusion

By its design and nature the game of cricket generates lots of data which lend itself to statistical and graphical analysis. In this paper a novel graphical display for comparing the performances of bowlers was considered. The graph was further extended to also be used for the comparison of batsmen and all-rounders. The graphs utilize ratios that are easily calculated using the batting and bowling records commonly provided in the printed and electronic media – see for example the Cricinfo website. A new performance criterion for batsmen, the survival rate, which is basically the balls per wicket ratio, was defined. To ensure uniformity between comparisons for bowlers and batsmen and to allow for the comparison of all-rounders, the use of standard terminology and notation was proposed.

Although the graphs were illustrated using batting and bowling records for cricketers that played in the IPL in 2008 under the T20 format of cricket, the graphs can be used for any format of cricket. This is not always true of other performance measures. For instance, the measures proposed above for assessing all-round performances are only applicable to limited-overs cricket.

The graphs can of course also be used in many other ways. For example, a graph can be constructed to compare the bowlers of two opposing teams before the start of a match. In the IPL the Rajasthan Royals played against the Chennai Super Kings in the final. In Table 3 the bowling records of the five prominent bowlers of each team are given. These records exclude the bowling figures in the final itself. In Figure 5 the corresponding ratios for the ten bowlers are plotted. Note that k was chosen as 6, so the horizontal axis represents the runs per over. Among the ten bowlers, the Chennai Super Kings had the two most expensive bowlers (L Balaji and Morkel) and the two bowlers with the highest balls per wicket and runs per wicket ratios (M Muralitharan and M Ntini). The Rajasthan Royals on the other hand, had the bowler with the lowest runs per over ratio (Sohail Tanvir) and the two bowlers with the lowest balls per wicket and runs per wicket ratios (Sohail Tanvir and SK Warne). Figure 5 therefore suggests that the Rajasthan Royals had the superior bowling attack of the two teams. This notion is supported by considering the combined ratios of the five bowlers of each team – see the codes “CSK” and “RR” in Figure 5.

Table 3. Bowling records of 10 bowlers who played in the final of the IPL in 2008*

Bowler	Code	Team†	Country‡	Balls	Runs	Wickets	RpB_6	BpW	RpW
L Balaji	LB	CSK	IND	174	244	11	8.41	15.82	22.18
M Gony	MG	CSK	IND	336	413	16	7.38	21.00	25.81
JA Morkel	JM	CSK	SA	264	374	15	8.50	17.60	24.93
M Muralitharan	MM	CSK	SL	324	365	9	6.76	36.00	40.56
M Ntini	MN	CSK	SA	186	221	7	7.13	26.57	31.57
Chennai↑	**CSK**			**1 284**	**1 617**	**58**	**7.56**	**22.14**	**27.88**
MM Patel	MP	RR	IND	318	406	14	7.66	22.71	29.00
Sohail Tanvir	ST	RR	PAK	223	226	21	6.08	10.62	10.76
SK Trivedi	TR	RR	IND	276	378	13	8.22	21.23	29.08
SK Warne	WA	RR	AUS	288	370	19	7.71	15.16	19.47
SR Watson	SW	RR	AUS	301	354	16	7.06	18.81	22.13
Rajasthan↑	**RR**			**1 406**	**1 734**	**83**	**7.40**	**16.94**	**20.89**

* Records do not include bowling figures from final itself.

† Teams:
CSK: Chennai Super Kings RR: Rajasthan Royals

‡ Countries:
AUS: Australia IND: India PAK: Pakistan SA: South Africa SL: Sri Lanka

↑ Only includes bowling records of the 5 prominent bowlers.

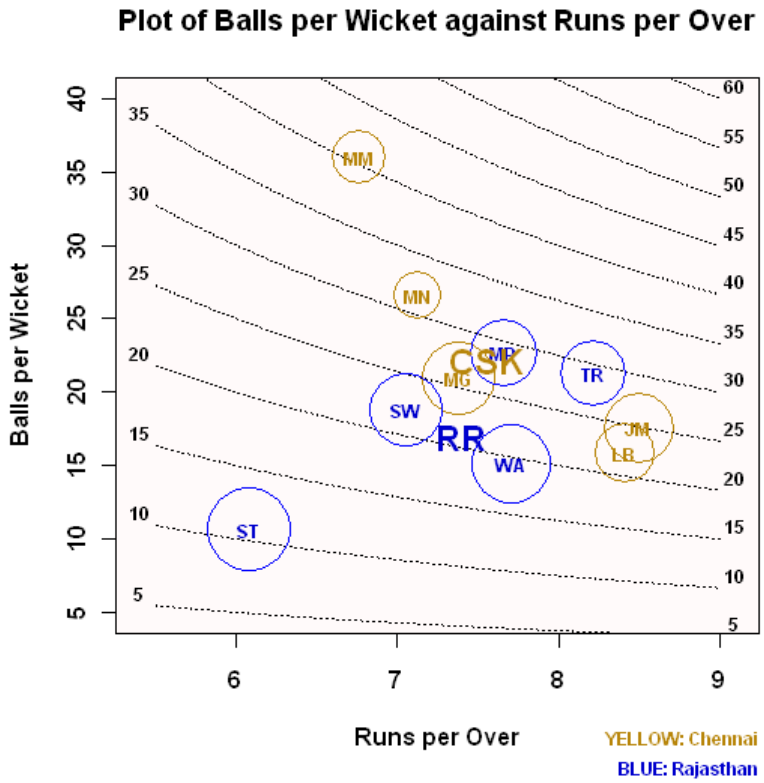

Figure 5. Comparison of the bowling performances of 10 bowlers who played in the final of the IPL in 2008

A cricketer's performance can also be traced over a series of matches. Sohail Tanvir's record for the IPL is given in Table 4. Note that he did not play in the first three matches of the Rajasthan Royals and also not in their ninth or fourteenth matches. In Figure 6 the ratios with respect to his cumulative bowling performance is

plotted, where the code "M5", as an example, indicates the fifth match of the Rajasthan Royals in the IPL. Because Sohail Tanvir did not take any wickets in his first match (his team's fourth match), BpW and RpW cannot be calculated after this match. He took his first wickets in his second match (his team's fifth match), so from this match onwards BpW and RpW can be calculated and plotted. In Sohail Tanvir's third match (his team's sixth match, played against the Chennai Super Kings) he obtained the best bowling figures in the IPL by taking 6 wickets for 14 runs in 4 overs. The effect of this performance is clear in Figure 6 in that the values of the ratios with respect to all three criteria deceased significantly. From then onwards there were no such drastic changes in the values of the ratios. Note however that his worst bowling performance (in terms of runs per 100 balls) occurred in the final of the IPL. Therefore the value for his runs per 100 balls ratio ended relatively high when compared to the values of his runs per 100 balls ratios over the IPL season.

Table 4. Cricket records of Sohail Tanvir for the IPL in 2008

		Match Records			Cumulative Records					
Code*	**Opponents†**	**Balls**	**Runs**	**Wickets**	**Balls**	**Runs**	**Wickets**	RpB_{100}	BpW	RpW
M1	DD	DID	NOT	PLAY	–	–	–	–	–	–
M2	KP	DID	NOT	PLAY	–	–	–	–	–	–
M3	DC	DID	NOT	PLAY	–	–	–	–	–	–
M4	BRC	24	31	0	24	31	0	129.17	–	–
M5	KKR	24	30	2	48	61	2	127.08	24.00	30.50
M6	CSK	24	14	6	72	75	8	104.17	9.00	9.38
M7	MI	19	13	0	91	88	8	96.70	11.38	11.00
M8	DC	24	39	0	115	127	8	110.43	14.38	15.88
M9	DD	DID	NOT	PLAY	115	127	8	110.43	14.38	15.88
M10	BRC	24	10	3	139	137	11	98.56	12.64	12.45
M11	KKR	24	26	3	163	163	14	100.00	11.64	11.64
M12	CSK	24	33	3	187	196	17	104.81	11.00	11.53
M13	MI	24	14	4	211	210	21	99.53	10.05	10.00
M14	KP	DID	NOT	PLAY	211	210	21	99.53	10.05	10.00
M15	DD	12	16	0	223	226	21	101.35	10.62	10.76
M16	CSK	24	40	1	247	266	22	107.69	11.23	12.09

* Indicates the number of each match of the Rajasthan Royals – they played in 14 league matches (M1 to M14), a semi-final (M15) and the final (M16).

† Opponents:

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
RR: Rajasthan Royals

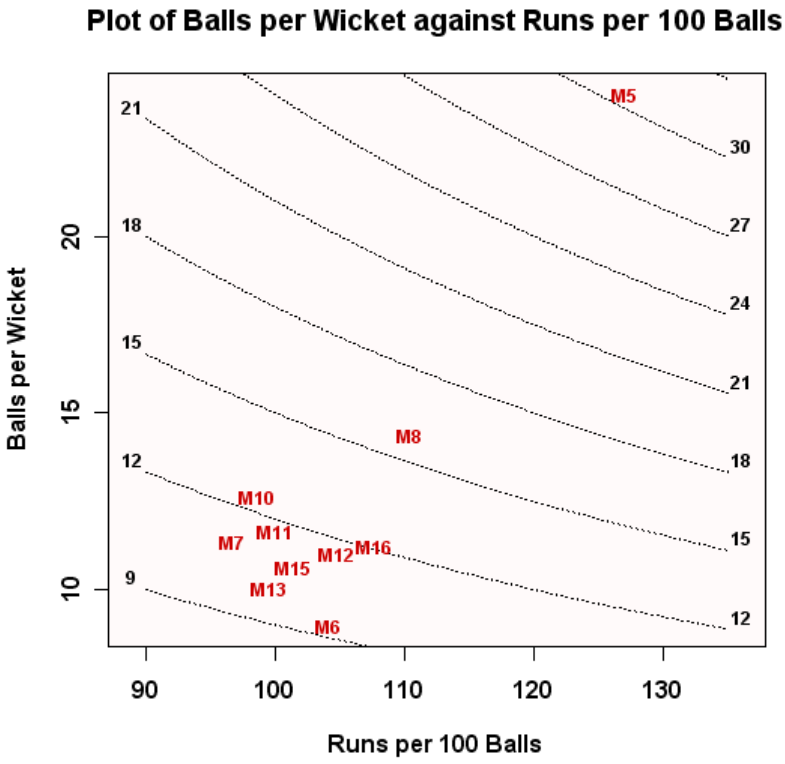

Figure 6. Tracing of the cumulative bowling performances of Sohail Tanvir in the IPL in 2008

Test cricket and one-day cricket present different challenges to cricketers. For example, in one-day cricket batsmen need to score runs at a faster rate than in Test cricket. Over the years some cricketers have become specialists in Test cricket or one-day cricket. T20 cricket is a very new format of cricket which requires different skills from the traditional formats of the game. It can be seen already from analyzing the IPL, that certain players, for example Marsh, Mishra, Morkel, Sohail Tanvir and Watson, may become specialist T20 players, whereas the style of play of others, like Mohammad Asif, Kallis and Styris, are maybe not suited to the game. As more T20 cricket is played around the world, it will become clearer who the specialist T20 cricketers are.

References

1. Bradman, D. G., *Farewell to Cricket*, Hodder and Stoughton, London, 1950, p.118.
2. Machale, D., What does "mean" mean? *The Mathematical Gazette*, 1990, **74**(469), 239–243.

3. Bolt, R. L., 19. Cricket averages. *The Mathematical Gazette*, 1958, **42**(340), 119–120.
4. Lemmer, H. H., The combined bowling rate as a measure of bowling performance in cricket. *South African Journal for Research in Sport, Physical Education and Recreation*, 2002, **24**(2), 37–44.
5. Lemmer, H. H., 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*, 2005, **27**(1), 75–87.
6. Lemmer, H. H., A measure of the current bowling performance in cricket. *South African Journal for Research in Sport, Physical Education and Recreation*, 2006, **28**(2), 91–103.
7. Kimber, A., A graphical display for comparing bowlers in cricket. *Teaching Statistics*, 1993, **15**(3), 84–86.
8. The Comprehensive R Archive Network. <http://cran.r-project.org/>.
9. Cricinfo: The home of cricket. <http://www.cricinfo.com/>.
10. Danaher, P. J., Estimating a cricketer's batting average using the product limit estimator. *New Zealand Statistician*, 1989, **24**(1), 2–5.
11. Kimber, A. C. and Hansford, A. R., A statistical analysis of batting in cricket. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 1993, **156**(3), 443–455.
12. Lemmer, H. H., Beraming van die kolfvermoë van Suid-Afrikaanse internasionale eendagkrieketspelers. *Suid-Afrikaanse Tydskrif vir Natuurwetenskap en Tegnologie*, 2001, **20**(2), 45–48.
13. Lemmer, H. H., Measures of batting performance in a short series of cricket matches. *South African Statistical Journal*, 2008, **42**(1), 65–87.
14. Barr, G. D. I. and van den Honert, R., Evaluating batsmen's scores in Test cricket. *South African Statistical Journal*, 1998, **32**(2), 169–183.
15. Lemmer, H. H. and Nel, J., A classification scheme for the consistency of batsmen. *South African Statistical Journal*, 2001, **35**(1), 147–173.
16. Lemmer, H. H., A measure for the batting performance of cricket players. *South African Journal for Research in Sport, Physical Education and Recreation*, 2004, **26**(1), 55–64.

17. Damodaran, U., Stochastic dominance and analysis of ODI batting performance: The Indian cricket team, 1989-2005. *Journal of Sports Science and Medicine*, 2006, **5**(4), 503–508.
18. Lewis, A. J., Towards fairer measures of player performance in one-day cricket. *Journal of the Operational Research Society*, 2005, **56**(7), 804–815.
19. Duckworth, F. C. and Lewis, A. J., A fair method for resetting the target in interrupted one-day cricket matches. *Journal of the Operational Research Society*, 1998, **49**(3), 220–227.
20. Duckworth, F. C. and Lewis, A. J., A successful operational research intervention in one-day cricket. *Journal of the Operational Research Society*, 2004, **55**(7), 749–759.
21. Beaudoin, D. and Swartz, T., The best batsmen and bowlers in one-day cricket. *South African Statistical Journal*, 2003, **37**(2), 203–222.
22. Johnston, M. I., Clarke, S. R. and Noble, D. H., Assessing player performance in one day cricket using dynamic programming. *Asia-Pacific Journal of Operational Research*, 1993, **10**(1), 45–55.
23. Clarke, S. R., Dynamic programming in one-day cricket – Optimal scoring rates, *Journal of the Operational Research Society*, 1988, **39**(4), 331–337.