



© 1997–2009, Millennium Mathematics Project, University of Cambridge.

Permission is granted to print and copy this page on paper for non-commercial use. For other uses, including electronic redistribution, please contact us.

---

March 2003

Features



## Howzat!

by Rob Eastaway



*Cricket is a sport that has always been popular with mathematicians, partly for the obvious reason that it is so full of numbers and statistics. For example, when you hear the number 501, you may think Levi jeans, but when cricket fans hear that number they also think Brian Lara. In 1994, Lara scored 501 runs in a single innings, the highest score ever made by an individual player in a professional cricket match. (In a sort of symmetry often loved by mathematicians, Lara was actually being sponsored by a different jeans company, Joe Bloggs, at the time!)*

## Ranking cricketers

Of course when there are lots of statistics, people like to use them to measure things, and the most tempting challenge of all is trying to measure who is the best player in the world.

## Howzat!

Since Victorian times, statisticians have largely relied on a very simple statistic – an *average* – to rank cricketers. For a batsman this average is extremely easy to calculate. Take the number of runs he has scored in his career and divide those runs by the number of times he has been out (that's caught out, or bowled, or lbw or all sorts of other more obscure dismissals).



However, even if you are not a cricket lover you may be able to think of some possible flaws in this measurement. In any team's innings in a cricket match, there is always at least one player who is "not out" at the end. Usually he's a pretty hopeless player who comes in at the end, and scores about 2 runs. But suppose by good fortune he happens to have a string of innings like this:

1\* 2\* 0\* 3\* 1\* 2\* 1

where \* indicates he was not out. His "average" at the end of this season is 10 runs divided by 1 out, i.e. an average of 10 runs per innings. This is pretty flattering given that he has never made it past 3!

There is another serious flaw when using averages to pick out the current best player. Imagine two players, Mark and Steve, whose career scores are like this:

Mark:	0	10	20	30	40	50	60	70	80	90	100
Steve:	100	90	80	70	60	50	40	30	20	10	0

A quick calculation will show that the mean score for each batsman is the same, i.e. 50. However, there is an obvious trend, in that Mark seems to be getting better, while Steve is getting worse. A raw average like a simple "mean" fails to identify trends.

Back in 1987, I was approached by the former England cricketer, Ted Dexter, to see if I could help him to devise a better method of rating cricketers statistically. As a young cricket and maths lover, what else could I say but "YES!"? I am still involved in running what are now known as the *PwC Ratings* to this day.

One obvious starting point in creating a cricket rating was to develop a statistic that would identify trends. Using the example above, we wanted a statistic that could spot that Mark should currently rate higher than Steve. A simple way of doing this is to just pick out the performances over a recent period of time, such as the last 12 months, or the last five matches. Indeed, this is how the world rankings in tennis are calculated. In the above example, using the last five matches would rate Mark as  $400/5 = 80$ , and Steve as  $100/5 = 20$ .



However, when developing a statistic that is to be presented to the general public, "mathematical sense" is not always the same as the public's "common sense", and you ignore the latter at your peril.

If in his next match, Mark scores 70 runs, this might suggest that his form is dipping slightly (because it is worse than his current rating of 80). However if we are looking at his last 5 performances, this score of 70 will displace the 60 he scored 5 matches ago, so his rating will actually go up. By the same process, if Steve now scores 30 - which is better than his current rating of 20 - his rating will actually go down because it displaces a 40. Mathematically this movement is perfectly understandable, but for the general public who haven't the time or inclination to go into the details of the calculations, such a counter-intuitive outcome leads to incomprehension or even ridicule (as the tennis authorities have learned in the past with their own rankings).

One of the best ways of avoiding the anomaly where a bad performance can result in your rating moving upward, or vice versa, is to use what is known as an *exponentially decaying average*. In this statistic, every score is considered in the calculation, but as you go back in time, each value is discounted by a certain percentage. If the decay rate is 4%, then we could calculate Mark's exponential average as follows:

Adjust his total runs over his 11 innings as follows:

$$100 + 90 \times 0.96 + 80 \times 0.96^2 + 70 \times 0.96^3 + 60 \times 0.96^4 + \dots + 0 \times 0.96^{10}.$$

That works out as 489.96.

Instead of dividing by 11 innings, you need to divide by

$$1 + 0.96 + 0.96^2 + \dots + 0.96^{10},$$

which comes to 9.04.

That gives a weighted average for Mark of  $489.96/9.04$ , or 54.2. That's the sort of result that we would expect, in that Mark has been performing better recently than his career average of 50 would suggest. And as a check that confirms that this method produces sensible results, if Mark scores 60 in his next innings his decayed average does indeed go up, if he scores 50 it goes down, and if he scores 54.2, it stays the same. (OK, so it's not actually physically possible to score 0.2 of a run, but you get the point.)

## Weighting problems



That decayed–average approach is the basic principle behind the cricket ratings system that we devised. However, there is a lot more to rating a cricketer than this, at least if you want a statistic that claims to be fair. If you are comparing players from one country against those from another, you have to allow for the fact that they are not playing in the same conditions as each other. (As protestors about school league tables might say, "everyone needs to be judged on a level playing field".) For example, if you are playing against Australia, who have some of the world's strongest bowlers, then scoring 50 runs is a much greater achievement than if you score 50 against Bangladesh, whose bowlers rate extremely low. A cricket rating should be able to reflect this. The rating system we developed makes a number of proportional adjustments to allow for different opposition strengths and high– or low–scoring matches.

The algorithm involved is rather too lengthy to go into here, but the result is a system that adjusts each player's individual performance in a match, and then updates the player's overall decayed average to produce a rating. There is a database that holds all previous ratings so that the fluctuations of the player's rating over his career can be plotted as a graph. This maps out his fluctuations in form in much the way that a graph can show the historic performance of your stockmarket investments (with similar surges and crashes).

Buried in the algorithm are some tricky little mathematical conundrums. For example, when updating a player's rating, the value of a batsman's performance needs to take account of the latest ratings of the opposing bowlers, while the bowler's performance needs to take account of the opposing batsmen's up–to–date ratings. There is a circularity here. Who should you rate first – the bowlers or the batsmen? We got around this problem by giving each player a "provisional" rating after the match, using that provisional rating in all the adjustment calculations, and then replacing the provisional figure with the updated figure. Although this introduces slight distortions, they are what might be described as "second order" errors and can be discounted.

## Testing the model

## Howzat!



The final and perhaps most delicate stage of producing a mathematical model of any kind, whether to rate cricketers or forecast the economy, is taking a look at the results and checking that they look "right". There are some real risks in this. The whole point of producing the model in the first place was that you didn't know what the "right" answer was. So if you now say "I don't like these results, let's change them so that they say *this*", you are in danger of negating all your objective work by superimposing your subjective judgments. And in any case, it's no good fudging the results so that they make today's results OK. It might be that all you are doing is papering over the cracks, so that the results the model produces tomorrow will throw up something absurd.

There are, however, tests you can do to make sure that the ranking you are producing does at least make sense. The following are some of the principles that we applied in setting up the cricket ratings, though they apply to all sorts of mathematical modelling beyond the world of cricket:

1. Use your common sense. If the computer says that the world number one batsman is Phil Tufnell (a man known for his fear of facing fast bowlers, not for his batting prowess), or that the UK economy next year is going to grow by 94%, you can be certain that there is an error somewhere either in your assumptions or in the execution of your mathematical formulae. Peculiar results are useful in pointing you to the areas that require closer inspection. (Though watch out – this doesn't necessarily work the other way around. Sometimes a faulty model produces sensible-looking results by fluke, so just because the results look OK doesn't mean that the model is OK.)
2. Test the model for sensitivity. If you have included a whole range of factors, it's worth seeing what happens if you tweak each of these factors a little, one at a time. Sometimes, adjusting a factor (like, say, adjusting the weighting given for opposition strength) has very little impact on the results. However, small adjustments in other factors might have a big impact. When a small adjustment to a factor leads to big changes in the output, this is clearly a highly sensitive area of the model, and it therefore requires the closest scrutiny.
3. Look at the long-term values of the output to see if any patterns emerge. One of the reassuring things about mathematical models is that however complex the algorithms, in the long term they tend to produce relatively predictable patterns. One particular pattern we looked for in the cricket ratings was

## Howzat!

an upward or downward drift in the numbers. On the assumption that over the last 30 years cricketers have remained at a similar level of ability, we expected the average point levels to be roughly the same for the years 1973 and 2003 (give or take a bit of "noise"). If there were signs of an upward drift, then this might mean players were getting better, but it was just as likely to mean that there was a feedback loop in the system that rated batsmen too highly, which in turn boosted the ratings of the opposing bowlers, which then boosted the batsmen who faced them later...and so on. Identifying the cause of drift is not easy, and to some extent this is where the intuition of the modeller comes in to make adjustments that are "reasonable".



If all this is making you think that producing a ranking table is as much an art as a science, then you would be right. Maths is a wonderful logical tool, but when it comes to measuring or representing the real world, it can't always guide you to the right answer on its own.

So when you next tune into a cricket match and hear the commentators say "and so, as world number one batsman Matthew Hayden\* strides out to the crease...", spare a thought for the mathematical model that's putting him there.

\* or Sachin Tendulkar, or Sanath Jayasuriya, or whoever it is this week.

---

## About this article



Rob Eastaway is an independent writer and lecturer, whose books include "Why do buses come in threes?" ([reviewed in Issue 10 of Plus](#)) and "How long is piece of string?" ([reviewed in Issue 21](#)). His article [Maths and magic](#), appeared in issue 14.

## Howzat!

You can find out more about Rob's various activities by visiting his [home page](#), and about the PwC cricket ratings by visiting [www.cricketratings.com](http://www.cricketratings.com).

All the pictures in this article were used by kind permission of [Cricket Ireland](#), a [Cricket Europe](#) site. Their [Photo Gallery](#) is well worth a visit.

---



*Plus* is part of the family of activities in the Millennium Mathematics Project, which also includes the [NRICH](#) and [MOTIVATE](#) sites.