METHODS FOR QUANTIFYING PERFORMANCES IN ONE-DAY CRICKET

Paul Allsopp and Stephen R. Clarke
School of Mathematical Sciences
Swinburne University
PO Box 218
Hawthorn
Victoria 3122, Australia

Abstract

Using the Duckworth and Lewis rain interruption rules for one-day cricket matches, a margin of victory in runs is created for the team batting second using the par, projected and new projected scores. The par score is shown to underestimate the margin of victory, whereas, estimates based on the projected and new scores are essentially equivalent to those obtained by the team batting first. If remaining resources for the team batting second are low, the results show that any differences in the margins of victory using either method are marginal. The resulting margin of victory is also used to model a team's rating and common home advantage in the Australian domestic one-day cricket competition. The differences in team ratings generated by both the projected and new projected scores are shown to be marginal and produce the same ranking order. The application of the projected and new projected scores showed that teams experience a common home advantage of nine and eleven runs, respectively, but these were not significant results. This is supported by application of binary logistic regression techniques. The overall ranking of teams produced by the model is also compared with the ranking based upon each team's mean margin of victory and shown to be in generally strong agreement.

1 Introduction

Duckworth and Lewis [6, 7, 9] have developed innovative rain interruption rules that are extensively used in one-day cricket matches. Their methods differ from previous approaches in that they take into account the available run-scoring resources (overs and wickets) the two competing teams have at their disposal. In summary, the more unutilised run-scoring resources a team has at their disposal at the point of interruption of a match the more runs they would be expected to make if the match was not interrupted.

This paper will adapt methods developed by Duckworth and Lewis [6, 9] and proposed by Clarke [2] and Allsopp and Clarke [1] to estimate a margin of victory (in runs) for the team batting second (i.e. Team 2), after they have gone on to win a match. The estimate will reflect the relative strength of Team 2 and provide a fair appraisal of their performance. We will then use adjusted Team 2 scores to investigate home advantage in the Australian domestic one-day cricket competition (1994–2000).

2 Calculating a margin of victory

2.1 The par score

Clarke [2] suggested that the methods developed by Duckworth and Lewis (D/L) to revise targets in matches interrupted by rain could be used to provide a margin of victory in runs. Using statistical data collected over a long period of time Duckworth and Lewis have developed a method that sets a revised target for Team 2 when overs in either innings have been lost due to a break in play. The target is revised in accordance with the available run-scoring resources the two teams have at their disposal. The adjusted targets ultimately reflect the relative difference in the resource availability of both teams.

In adapting the D/L method to reflect the relative strength of Team 2, when they have gone on to win a match, we treat the completion of Team 2's innings as a break in play. Team 2, in winning the match, has subsequently used up less of their available run-scoring resources in surpassing the target set (unless they win on the last ball). If we denote the revised target for Team 2 by T, Team 1's total score by S, the run-scoring resource percentage remaining by R and the run-scoring resource percentage available to Team 2 by $R2$, then T can be calculated by:

- Scaling Team 1's score downwards in the ratio $R2$ to 100. This is the score to tie.
- Adding one to give the target.

Knowing the number of overs left and the number of wickets Team 2 has lost, Duckworth and Lewis [6, 7] have prepared detailed tables from which the appropriate $R2$ values can be determined. Reading the tables directly provides the resource percentage remaining for Team 2, denoted by R. It follows, $R2 = 100 - R$. As defined by Duckworth and Lewis [6, 7, 9], it follows:

$$T = S\frac{R2}{100} + 1 = S\frac{100 - R}{100} + 1. \tag{1}$$

If a match is abandoned during the second innings, $(T-1)$ is defined by Duckworth and Lewis as the par score, or the score that Team 2 will need to have achieved in order to tie the match at this point. If Team 2 is ahead of its target, Duckworth and Lewis [6] quantify the difference between the current and par scores as Team 2's margin of victory. However, if Team 2 is behind the target set at this point the par score is denoted by Team 1's score and the difference between the two scores will be Team 1's margin of victory. At the completion of a match, the par score represents the score that Team 2 will need to have compiled in order to achieve a tied result at the point their innings is completed. If Team 2 wins the match they are obviously ahead of their target and the subsequent difference between the actual and par scores is defined as Team 2's margin of victory. However, if Team 2 is behind the set target at this point, the par score is defined as Team 1's total score and the difference between the two scores in this case will be Team 1's margin of victory.

2.2 The projected and new projected scores

Using the par score to determine the margin of victory gives some indication of how well Team 2 has performed but it does not tell the whole story since we don't know how many more runs Team 2 could go on to make if they batted out their 50 overs. If, for example, Team 2 wins we can only be certain of how far Team 2 is ahead of its target at the completion of their innings irrespective of how many run-scoring resources Team 2 has at its disposal. We will demonstrate that the par score is not a fair indication of how well Team 2 has performed because the margins of victory that are generated will not be equivalent to those obtained by Team 1. We propose that an estimate of Team 2's projected 50-over score will form the basis of a more accurate measure of the margin of victory. This estimate will be based on two methods, namely the projected score and the new projected score. The projected score is the sum of Team 2's current score at the completion of their innings and an estimate of the number of runs they will make in the remaining overs. This estimate is a percentage of 225 runs, which Duckworth and Lewis define as the average score compiled by teams in a 50-over innings. Alternatively, the new projected score assumes Team 2's final score is the par score. This score can then be used to determine the equivalent score Team 1 needed to have achieved in order to tie the match at this point. This in effect provides a measure of how well Team 2 has performed. In both cases the resultant margin of victory is the difference between Team 1's score and Team 2's adjusted score.

If X represents Team 2's current score, to estimate Team 2's projected 50-over score P, we have:

$$P = X + 225\frac{100 - R2}{100} = X + 2.25(100 - R2) = X + 2.25R. \tag{2}$$

If we define Team 2's new projected score as N, then

$$N = \frac{100X}{R2} = \frac{100X}{100 - R}. \tag{3}$$

3 Analysis of the Australian domestic one-day cricket competition (1994–2000)

Tables 1 and 2 provide a summary of results from the Australian domestic one-day cricket competition (1994–2000) which includes results from 117 completed matches. The results are for winning teams only. Figure 1 provides a series of boxplots showing the distribution of the winning scores. The boxplots suggest that, on average, Team 1 posted higher winning scores than Team 2. This is expected, since Team 2, in winning a match, has its innings truncated as soon as they pass Team 1's score. Notably, the winning scores of Team 2 are more variable than those generated by Team 1. The par score also represents a truncated score and is, on average, lower than all listed scores, however, because the par score can result from a relatively wide range of overs, it is more variable. The projected score, on average, is higher than the new projected score, but due to the presence of outliers the new projected score is the more variable. The outliers for the new projected score result from matches in which Team 2 quickly passed Team 1's score with many unutilised run-scoring resources at their disposal.

In comparing the distribution of winning scores the normality assumption (Anderson–Darling test) holds for Team 1 and Team 2's actual winning scores and for both the par and projected scores, but is violated for the new projected score. Using a two-sample t-test the analysis clearly suggests that, on average, Team 1's score is significantly higher than both Team 2's actual score ($p = 0.000$) and the par score ($p = 0.000$). This results from the fact that Team 1, in winning their matches, exhaust available run-scoring resources and so maximise their return. However, Team 2, in winning always has unutilised run-scoring resources at their disposal (unless they win off the last ball) and so is not able to maximise their run scoring potential. Using the non-parametric Mann–Whitney test to compare distribution of scores, both the projected and new projected scores are not significantly different from Team 1's winning score ($p = 0.390$ and $p = 0.535$, respectively). The new projected score is also not significantly different from the projected score ($p = 0.663$).

	Team 1	*Team 2*	*Par score*	*Projected score*	*New projected score*
Mean	248.4	205.2	170.4	243.1	249.9
Median	242.0	208.0	178.3	237.0	238.4
Standard deviation	32.8	38.0	48.9	33.5	47.2

Table 1: Summary of results from the Australian domestic one-day cricket competition (1994–2000).

Tables 3 and 4 provide a respective summary and analysis of the winning margins of victory generated by Teams 1 and 2. Figure 2 provides a series of boxplots showing the distribution of the margins of victory. With reference to the boxplots, the margins of victory (in runs) generated by Team 2's actual score are inconsequential since Team 2's innings is truncated once Team 1's score is surpassed. Notably, Team 1, on average, generated the higher margins of victory, which were also the most variable. The margins of victory generated by the par, projected and new projected scores, on average, were similar, with the new projected score clearly the more variable. Notably, application of the new projected score to generate a margin of victory has produced a relatively high number of outliers. This arises because the method predicts relatively high scores for Team 2 when they have won a match with a high proportion of unutilised run-scoring resources still at their disposal.

	Comparison of scores	*Test*	*p-value*
H_1	Team 1 > Team 2 (actual)	Two-sample t-test	0.000
H_1	Team 1 > Par	Two-sample t-test	0.000
H_0	Team 1 = Projected	Two-sample t-test	0.390
H_0	Team 1 = New projected	Mann–Whitney	0.535
H_0	Projected = New projected	Mann–Whitney	0.663

Table 2: Competition analysis.

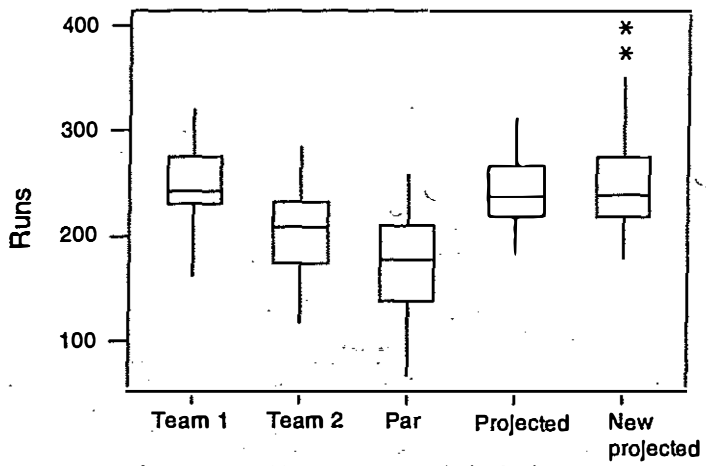

Figure 1: Distribution of winning scores.

Since the normality assumption is violated for all distributions of the margins of victory (Anderson–Darling test), the Mann–Whitney test is used to compare the distributions. The analysis suggests that the margins of victory generated by both the projected score ($p = 0.072$) and the new projected score ($p = 0.190$) are not significantly different from the margins of victory obtained by Team 1. However, the par score generated margins of victory that were significantly less than those generated by Team 1 ($p = 0.005$) and so underestimated the margins estimated by the projected and new projected scores. Both the projected and new projected scores generated margins of victory that were not significantly different from those resulting from adopting the par score ($p = 0.600$ and $p = 0.245$ respectively). Notably, there is no significant difference between the margins of victory generated by the projected and new projected scores ($p = 0.566$).

In relatively few instances the margin of victory generated by the par score exceeds the margin of victory generated by the projected score. This anomaly arises whenever $R(S - 225) > 0$. Since in all cases $R > 0$ (i.e. after at least one ball has been bowled), this situation only arises when Team 2 has gone on to win a match after Team 1 has posted a score in excess of 225. This suggests that Team 1, in losing a match, has performed better than average.

	Team 1	*Team 2*	*Par score*	*Projected score*	*New projected score*
Mean	53.4	2.1	34.8	40.0	46.9
Median	41.0	2.0	30.3	31.7	33.0
Standard deviation	43.2	1.2	21.5	29.1	43.1

Table 3: Margins of victory results.

Figures 3, 4 and 5 represent plots of the differences between the margins of victory generated by

	Comparison of margins of victory	*Test*	*p-value*
H_1	Team 1 > Par	Mann–Whitney	0.005
H_0	Team 1 = Projected	Mann–Whitney	0.072
H_0	Team 1 = New projected	Mann–Whitney	0.190
H_0	Par = Projected	Mann–Whitney	0.600
H_0	Par = New projected	Mann–Whitney	0.245
H_0	Projected = New projected	Mann–Whitney	0.566

Table 4: Analysis of margins of victory.

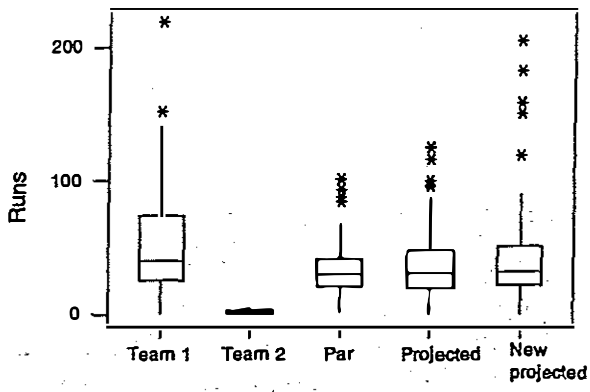

Figure 2: Distributions of the winning margins of victory.

the par, projected and new projected scores against the number of overs remaining. The plots suggest that when the number of overs remaining is relatively small the resulting differences in the margins of victory generated by each representation of Team 2's winning score are also relatively small. However, this difference incr

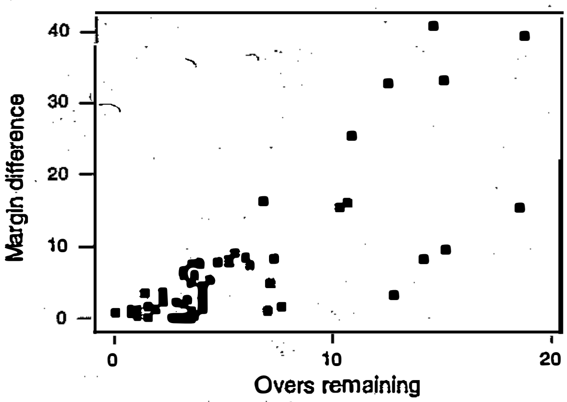

Figure 3: Plot of the differences in margins between the par and projected scores.

Table 5 provides a summary of the correlation coefficients between the margins of victory generated by each representation of Team 2's score. The results suggest that for all winning scores there is a strong positive correlation of 0.949 ($p = 0.000$) between the margins generated by the projected and new projected scores. However there is evidence of only a moderate positive correlation between the margins generated by the par and projected scores and the par and new projected scores (coefficients are 0.636 ($p = 0.000$) and 0.465 ($p = 0.000$), respectively). These observations are deceptive since as is

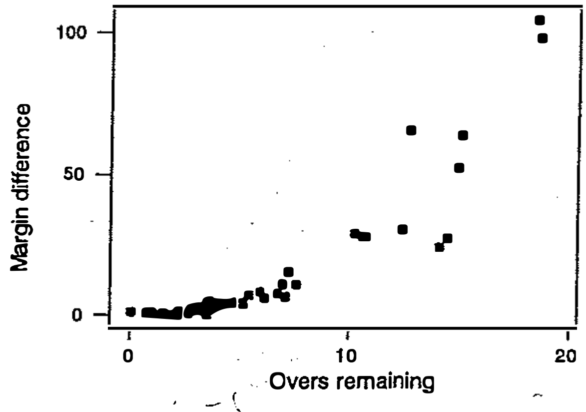

Figure 4: Plot of the differences in margins between the par and new projected scores.

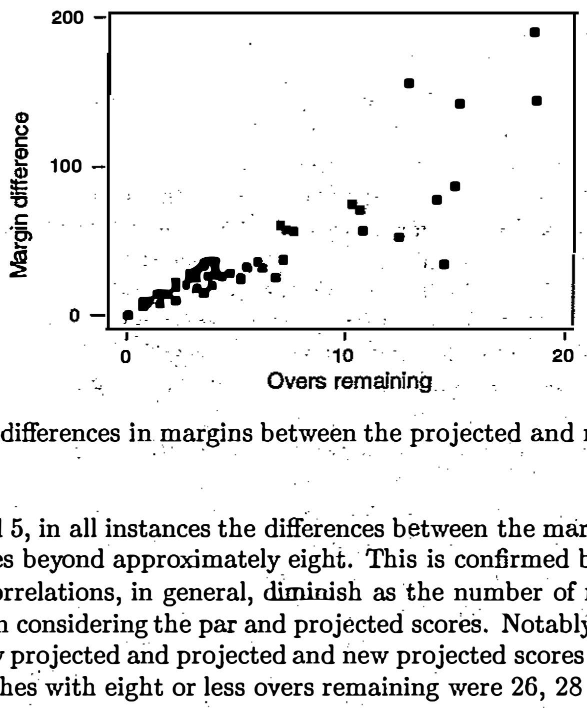

Figure 5: Plot of the differences in margins between the projected and new projected scores.

clear from Figures 3, 4 and 5, in all instances the differences between the margins inflate as the number of overs remaining increases beyond approximately eight. This is confirmed by Table 5, which suggests that the strength of the correlations, in general, diminish as the number of remaining overs increases. This is most apparent when considering the par and projected scores. Notably, when comparing the par and projected, par and new projected and projected and new projected scores the mean difference in the margins of victory in matches with eight or less overs remaining were 26, 28 and 29 runs, respectively. The mean differences increased to 70, 92 and 120 runs, respectively, when the number of overs remaining was nine or more.

Comparisons		*Correlation coefficient*		
		All winning scores	*Less than nine overs remaining*	*More than nine overs remaining*
Par score	Projected score	0.310	0.926	0.730
Par score	New projected score	0.105	0.995	0.960
New projected score	Projected score	0.959	0.946	0.882

Table 5: Summary of correlation coefficients.

4 Analysing home advantage in the Australian domestic one-day competition (1994–1999)

The Australian domestic one-day cricket competition is currently referred to as the Mercantile Mutual Cup and is played between teams representing the six States of Australia and the Australian Capital Territory. The competition is a round-robin tournament with teams gaining two points for a win or one point each for a tie or a "no result". The top four teams at the end of the round robin play off in two semi-finals, with the winners playing each other in the final.

Home advantage is formally defined as the expected difference in score in a game played between two teams (on the home ground of one of the teams) minus the expected difference in score in a game played between the same two teams on a neutral ground. In the context of one-day cricket, home advantage represents the margin of victory in a game played between team i and team j (on the home ground of team i) minus the margin of victory in a game played between team i and team j (on the home ground of team j).

Using techniques adopted by Stefani and Clarke [10], Harville and Smith [8] and Clarke and Norman [4] the winning margin w_{ij} in a match between team i and team j played at the home ground of team i is modelled as:

$$w_{ij} = (u_i + h) - u_j + \epsilon_{ij} = u_i - u_j + h + \epsilon_{ij} \tag{4}$$

where u_i is a measure of the relative ability of team i, h is a measure of the common home advantage and ϵ_{ij} is a zero-mean random error.

A least squares regression model has been fitted to the margins of victory (generated by both the projected and new projected scores) to quantify (a) a team's rating, and (b) any common home advantage. It is assumed that a team's average rating is 100. Table 5 provides a summary of the ratings for each team together with the mean margin of victory over the period 1994 to 2000. Notably, the choice of whether to choose the projected or new projected score to calculate the margin of victory has (a) produced similar ratings and (b) preserved the same ranking order for each team. It is also notable that the teams, on average, did not experience a significant common home advantage under either method, with the advantages generated by the projected and new projected scores being only nine runs ($p = 0.088$) and ten runs ($p = 0.057$), respectively.

In considering the outcome only of each match (i.e. home win/home loss and away win/away loss), we have the home and away teams winning 54% and 46% of matches, respectively. In applying a binary logistic regression model, there is some evidence that the home team experiences an advantage. However, any advantage is not statistically significant ($z = 1.18$, $p = 0.238$), with the odds of winning away being about 1.4 times the odds of winning at home.

Using the mean margin of victory (generated by both the projected and new projected scores) to rank each team shows generally strong agreement with the ranking produced by the model estimates. Home advantage based on estimates generated by the mean margin of victory for each team (i.e. four and six runs for the respective projected and new projected scores) showed some agreement with the model estimates.

5 Conclusions

The use of the D/L method to deal with one-day cricket matches interrupted by rain is well documented and has been used in a number of competitions. The method can also be effectively adapted to provide a relative measure of how well the team batting second has performed by generating a margin of victory in runs equivalent to the team batting first. The margin of victory is a more sensitive measure of the strength of a win or loss.

The par score provides some measure of Team 2's relative performance. However, it does not generate a 50-over based margin of victory. Consequently, use of the par score tends to underestimate Team 2's margin of victory. Using both the projected and new projected scores provides a fairer appraisal of

Team	*Projected score*		*New projected score*	
	Rating	*Mean margin of victory*	*Rating*	*Mean margin of victory*
Western Australia	128	30	132	41
Queensland	124	24	124	25
New South Wales	116	14	118	15
Tasmania	93	−10	97	−6
South Australia	92	−2	89	−7
Victoria	83	0	79	0
ACT	65	−27	61	−26
Common home advantage	9	4	11	6

Table 6: Rating of teams in the 1994–1999 domestic one-day competition.

Team 2's relative performance. In each case the scores generate a margin of victory essentially equivalent to that obtained by Team 1. This suggests that the margins of victory generated by the projected and new projected scores provide a more accurate measure of a team's performance.

In matches won by Team 2, when the number of remaining overs was relatively small, any difference in the margins of victory generated by the par, projected and new projected scores was marginal. This suggests that when the number of overs is low (less than eight) the margin of victory generated by either method will in effect be equivalent to the margin of victory obtained by Team 1. However, as the number of remaining overs increases these differences become statistically significant and so it is more appropriate to use either the projected or new projected scores to generate a margin of victory in these cases.

Using the margin of victory generated by the projected and new projected scores to model team performance in the Australian domestic one-day competition (1994–2000) showed that the team-rating estimates were similar for both methods and the teams were ranked in the same order. Using each team's mean margin of victory to rank the teams showed general agreement with the rankings obtained by the model.

Based on scores estimated by the projected and new projected methods the teams, on average, experienced a common home advantage of eight and nine runs respectively. These results were not statistically significant. Home advantage based on estimates generated by the mean margin of victory showed some agreement with the model estimates. The application of binary logistic regression techniques also support the notion that teams on average did not experience a significant home advantage, with the odds of winning away estimated to be about 1.4 times the odds of winning at home.

References

[1] P. Allsopp and S. R. Clarke, "Fair measures of performance in one-day cricket", in *Mathematics 2000 Festival*, M. Bulmer *et al.* (editors), Melbourne, Victoria, Melbourne University (2000), 1–8.

[2] S. R. Clarke, *Performance Modelling in Sport*, unpublished PhD thesis (1998), Chapter 2.

[3] S. R. Clarke, "Test Statistics", in *Statistics in Sport*, J. Bennett (editor), Arnold, London (1998), 83–101.

[4] S. R. Clarke and J. M. Norman, "Home ground advantage of individual clubs in English soccer", *The Statist.*, 44(4) (1995), 509–521.

[5] S. R. Clarke and P. Allsopp, "Fair measures of performance: The World Cup of Cricket", submitted.

[6] F. C. Duckworth and A. J. Lewis, "A fair method for resetting the target in interrupted one-day cricket matches", *J. Oper. Res. Soc.*, **49** (1998), 220–227.

[7] F. Duckworth and T. Lewis, *Your complete guide to the Duckworth and Lewis method for resetting targets in one-day cricket*, University of the West of England, Bristol (1999).

[8] D. A. Harville and M. H. Smith, "The home-court advantage: How large is it, and does it vary from team to team?", *Amer. Statist.*, 48(1) (1994), 22–28.

[9] T. Lewis and Duckworth, "A fair method for resetting the target in interrupted one-day cricket matches", in *Third Conference on Mathematics and Computers in Sport*, N. de Mestre (editor), Bond University, Queensland, Australia (1996), 51–68.

[10] R. T. Stefani and S. R. Clarke, "Predictions and home advantage for Australian Rules football", *J. Appl. Statist.*, **19**(2) (1992), 251–261.