THE CHARACTERISTICS OF VARIOUS MEN'S TENNIS DOUBLES SCORING SYSTEMS

Brown, Alan[1], Barnett, Tristan[2], **Pollard, Geoff**[2], Lisle, Ian[3] and Pollard, Graham[3]

[1] Faculty of Engineering and Industrial Sciences, Swinburne University of Technology, Melbourne, Australia

[2] Faculty of Life and Social Sciences, Swinburne University of Technology, Melbourne, Australia

3 Faculty of Information Sciences and Engineering, University of Canberra, Australia

Submitted for Review ??/??/08; Accepted ??/??/08 (editors will insert the dates)

Re-submitted ??/??/08; Accepted ??/??/08 (include if re-submission is required)

Abstract. In recent times a range of best of five and best of three sets tennis scoring systems has been used for elite men's doubles events. These scoring system structures include advantage sets, tie-break sets, match tie-breaks, tie-break games, advantage games and no-ad games. Several tournament organizers, tennis administrators, players and ATP Tour representatives are interested in comparing the characteristics of these various scoring systems. These characteristics include such things as the likelihood of each pair winning, and the mean, the variance, and the 'tails' of the distribution of the number of points played under the various systems. In this paper these characteristics and the distribution of the number of points in a match are determined for these various doubles scoring systems at parameter values that are relevant for elite men's doubles. Advantage games and no-ad games, both approved alternatives within the rules of tennis, are considered, as is the '50-40' game.

Keywords: new scoring systems for men's tennis doubles, no-ad and '50-40' games of tennis.

INTRODUCTION

In a match of tennis, points are played to determine the winner of a game, games are played to determine the winner of a set, and sets are played to determine the winner of a match. Traditionally, a game is the best of six points, but if the score reaches 3-3, play continues until one player leads by two points. A traditional advantage set is the best of 10 games, but if the games score reaches 5-5, play continues until one player leads by two games. A match is the best of 5 sets, or the best of 3 sets.

This scoring system survived unchanged throughout the amateur era until 1968 when tennis was opened up to professional players, and tournaments became major television events. The first significant change was the introduction of the tie-break game at six games all in all sets (except in some cases the last set) to determine the winner of a set. Also, all ATP Tour tournaments gradually reduced men's doubles and then men's singles to best of 3 sets. More recently, men's and women's doubles on the ATP and WTA Tours have introduced sudden death or no-advantage scoring at deuce so that whoever wins the next point (instead of leading by two points) wins the game. They have also replaced the last set with an extended tie-break game, known as a match tie-break (10 points).

These and other recent modifications to the traditional scoring system were primarily designed to address the fact that the requirement to lead by two points to win a game, or lead by two games to win a set, produces a match of unknown and quite variable length that leads to considerable scheduling difficulties for tournament organizers and television coverage in particular. It was also hoped that these changes would encourage top singles players to compete in the doubles as well, but the principal objective, however, was to decrease both the mean and the variance of the length of a tennis match.

Very little research has been carried out on the effect of the various scoring system options on the mean and standard deviation of the length of tennis doubles matches. Pollard et al. (2007) studied the effect of the most recent changes to the scoring system used for men's doubles for ATP Tour events.

In this paper the traditional and all of the commonly used scoring systems in professional tennis are analyzed as well as other possible systems in order to advance the prime objective of reducing the length and variance of a tennis match without significantly affecting the overall probability that the better pair would have won if the traditional scoring system was still in place.

The following eight scoring systems have been identified and analysed using advantage games. These systems are called systems 1(a) to 8(a).

1. Best of 5 advantage sets. The traditional men's scoring system, allowable under the International Tennis Federation's 'Rules of Tennis 2008', but no longer in use.
2. Best of 5 sets, first four tie-break and fifth set an advantage set. Used in men's singles Grand Slams (except the US Open) and Davis Cup singles and doubles, as well as men's doubles at Wimbledon.
3. Best of 5 tie-break sets. Used in men's singles at US Open and final of Tennis Masters Cup.
4. Best of 5 sets, first four tie-break sets and fifth set a match tie-break (10 Points). Not currently used but a possibility under the present rules of tennis, and consistent with system 8. below.
5. Best of 3 advantage sets. The traditional scoring system for women and shorter men's matches, allowable under the Rules of Tennis 2008, but not presently in use.
6. Best of 3 sets, first two tie-break and third set an advantage set. Used in women's matches at Grand Slams (except the US Open) and at the Federation Cup.
7. Best of 3 tie-break sets. Used in doubles at most Grand Slams (not Wimbledon), and also used for women's singles at US Open and for men's and women's singles on ATP and WTA Tours.
8. Best of 3 sets, first two tie-break sets and third set a match tie-break (10 Points). This is used for mixed doubles at most Grand Slams, and for doubles on the ATP and WTA Tours.

Given that the ATP and WTA have now introduced no-ad scoring at deuce for doubles on their Tours, the above analysis was repeated using no-ad games. These scoring systems are called systems 1(b) to 8(b).

Also considered is the '50-40' game scoring system (Pollard & Noble, 2004), whereby the server is required to reach 50 (i.e. one more point than 40) before the receiver reaches 40 in order to win the game, whilst the receiver wins the game by reaching 40. The above eight systems using this '50-40' game are called systems 1(c) to 8(c). Thus, in total, 24 scoring systems are considered in this paper.

In tennis the server has an advantage over the receiver, and therefore a greater than 50% chance of winning the point. This chance is generally greater for men than for women, greater on grass than on clay, and greater in doubles than in singles. The focus of this research is on professional men's doubles, where 'long matches' can be an issue and where alternative systems are being trialed. For this reason, the analysis in this paper considers situations where the probability of the server winning the point varies from near 0 60 to near 0.75. Given that the relevant point-probability values for the servers depend on a range of things such as serving form on the day, the receivers' form on the day, the rankings of each pair, the court surface, the weather conditions, etc., it was considered that this range from 0.6 to 0.75 represented an appropriate range for covering men's professional doubles matches, at the present time and into the future.

For this analysis, the length of a match is measured by the number of points played under the various scoring systems. Obviously the actual time to play the match (in hours and minutes) also depends on other factors such as the court surface and the style of play. For example, the average time taken to play a point on a clay surface is typically longer than the time taken on a grass surface. Also, some players are known to play 'longer' points than others, and/or take more time between points. These factors have not been considered in this paper, but could be 'factored-in' by using additional information.

Following the changes in 2006 to the best of 3 sets scoring system used for men's doubles in a range of professional tournaments, there are now several scoring systems commonly used for professional men's doubles. In this paper the characteristics of the various scoring systems presently used in men's doubles or ones currently available under the Rules of Tennis 2008, are determined, and displayed in a quantitative manner. The various scoring systems are then compared with the view to determining how well they achieve the types of objectives mentioned above. The primary objective of this paper was to evaluate the characteristics of the present doubles systems, used or allowable under the rules. Given the research nature of this publication, however, it was considered appropriate to consider the merits of a new game scoring system, not in the Rules of Tennis 2008, but one that looked particularly interesting with regard to its possible use in men's doubles. This new game scoring system, the '50-40' game, is considered in the next paragraph. It is noted that its possible use in best of 5 sets matches has not previously been studied.

The usefulness of the '50-40' game, particularly in men's doubles, has been noted (Pollard & Noble, 2004). They noted that, for the p-values relevant to professional men's doubles matches (values typically greater than 0.60), 'the longest [best of 3 tie-break sets] matches (as measured by the 98% point in the distribution of the number of points played) can be reduced by about 30 points [by using '50-40' games instead of no-ad games]'. They also noted that for such p-values 'the probability player A wins [a best of 3 tie-break sets match] using '50-40' games is comparable to when using no-ad games, even though about 20 less points are required on average'. Note that the seventh point in the no-ad game creates a lack of symmetry with respect to serving to the right half or the left half of the court, and also creates a potential for unfairness, but these two issues are removed by using the '50-40' game. Thus, it would appear that the '50-40' game has some merit relative to the no-ad game for men's doubles.

What are the desirable (statistical) characteristics of a good tennis scoring system? The 'three-nesting' aspect of tennis (points, games and sets) is taken as a given or fixed part of tennis scoring systems, as it allows either player to overcome a period of poor play. Games and sets might be made longer or shorter than at present if there is an advantage in doing so. It is fundamental, however, that a scoring system should have an appropriate average number of points played, and an appropriate value for the probability that the better player wins. Also, the standard deviation of the number of points played should not be too large, so that matches have a reasonably predictable duration. Strongly positively skewed distributions of duration are to be avoided as, under such scoring systems, very long matches can result, and this can delay the matches that follow, can lead to unfairness in the tournament setting, and is obviously an issue for television. Scoring systems with good efficiency at correctly identifying the better player are preferable to ones with not-so-good efficiency. Given that a men's doubles tournament involves matches between very strong servers as well as matches between weaker servers, what is needed is a scoring system that works satisfactorily at both ends of this spectrum. This is an important consideration in selecting a tournament scoring system.

Given that the above characteristics need to be considered before adopting a tournament scoring system, compromises may need to be made between them when choosing one particular scoring system over another.

ANALYSIS OF VARIOUS SCORING SYSTEMS

For the 24 scoring systems under consideration, the probability that the better pair wins the match, and the mean and higher moments of the number of points played, were evaluated using recursion methods. The details of this (exact) recursion method are omitted here, but are available (Brown et al., 2008).

The process used to estimate the distribution of the number of points in a match required some care since, for a best of 3 sets match, it may be bi-modal, depending upon the design of the third set. Furthermore, a distribution is not uniquely determined by its moments. To overcome these difficulties, the distributions for matches requiring 2 sets or 3 sets to complete were estimated separately. This was done using the Normal Power approximation, and the first four moments of each distribution. The Normal Power approximation uses a basic assumption that the distribution is uni-modal, and it would be inappropriate to use it in conjunction with the statistics for an overall match. The probability weighted sum of the two distributions (or three in best of 5 sets matches) was used to estimate the distribution of the number of points in the overall match, from which the 98% point was obtained by interpolation. This point in the distribution is used as a measure of 'long' matches. This method was used in earlier work (Pollard et al., 2007). Further, the results agreed with those in the paper by Pollard & Noble (2004), who used simulations of 1,000,000 matches. With the exception of some cases of the 98% statistic, all statistics agreed with simulations of 4,000,000 carried out by the authors in 2008, and with other exact results (Pollard, 1983). If a difference of more than 2 points existed in the two estimated 98% statistics, the simulated value was reported in the tables.

We note here that two scoring systems can be compared for their efficiencies (at correctly identifying the better pair). Thus, given scoring system 1 and scoring system 2 with the same expected number of points played in a match, scoring system 1 is said to be more efficient than scoring system 2 if scoring system 1 has a higher value for the probability that the better pair wins the match. The efficiencies of scoring systems with differing values for the expected number of points played can also be evaluated, as in the elegant paper by Miles (1984). It is noted here that scoring systems with high efficiencies (i.e. efficiencies close to 1) typically have particularly large standard deviations, and hence are not appropriate in the sporting context.

The best of 5 sets systems 1, 2, 3 and 4 above using (a) advantage games, and (b) no-ad games, being approved tennis scoring systems, are considered firstly, and then these systems using (c) '50-40' games are considered. The following questions were considered. What are the characteristics of these various scoring systems? What are the differences between them? Do some of them have good/not-so-good characteristics?

Five characteristics are sufficient to make relevant comparisons of the above scoring systems. They are P(A wins) where A is the better pair, the mean and the standard deviation of the number of points played, the efficiency of the scoring system, and the 98% point of the cumulative distribution of the number of points.

These five statistical measures were evaluated for a range of p_a (the probability pair A wins a point when serving) and p_b (the probability pair B wins a point when serving) values. This range covered p-values of 0.60, 0.65, 0.70 and 0.75, with pair-deviations of 0, plus and minus 0.02, and plus and minus 0.04. A close inspection of all of these results indicated that it was sufficient to report on just the results for p-values near 0.60 and for p-values near 0.75, as the results for the intermediary parameter values lay between the results at these values. Thus, the statistical measures reported in table 1 are for matches at the 'strong-serving' end ($p_a = 0.77$, $p_b = 0.73$), and at the 'weaker-serving' end ($p_a = 0.62$, $p_b = 0.58$) of the above range.

It is useful to introduce a definition at this stage. We say a scoring system X is 'a better scoring system' than scoring system Y if it has an equal or larger P(A wins) value, an equal or smaller mean, an equal or smaller standard deviation, an equal or lower 98% point, and an equal or greater efficiency (provided there is at least one inequality). It is noted that if scoring system X is better than scoring system Y, and if scoring system Y is better than scoring system Z, then scoring system X must be better than scoring system Z.

A comparison of the best of 5 sets systems 1, 2, 3 and 4, using advantage and no-ad games

Firstly, comparing scoring systems 1(a) and 1(b), the means of systems 1(a) and 1(b) for a match between strong servers are very large (484.3 and 366.7), as are the standard deviations (218.2 and 149.6), and the 98% points (1049 and 755). Even for matches between weaker servers, the standard deviations and the 98% points are large ((73.8 and 61.6) and (442 and 377) respectively). It is these characteristics that caused system 1(a), that survived the amateur era, to be replaced by others in the professional era.

Secondly, comparing scoring systems 2(a) and 2(b), the standard deviations of both systems 2(a) and 2(b) for a match between strong servers, are large (99.5 and 76.7), as are the 98% points (582 and 467). Even for a match between weaker servers, these statistics are still quite large ((63.6 and 54.2) and (395 and 341)). Given the size of these characteristics, systems 2(a) and 2(b) (although accepted in Grand Slam singles) are only really applicable to the finals of important events, where the winners are not required to play another match on the following day, and in situations in which there is no match scheduled to follow. For matches between strong servers, scoring system 2(b) is very close indeed to being a better system than system 2(a).

Thirdly, comparing scoring systems 3(a) and 3(b), there is not a huge difference in the means, in the standard deviations and in the 98% points for the matches between strong servers and those between weaker servers ((272.0 and 248.9), (261.0 and 229.9); (60.7 and 55.7), (61.6 and 52.7); and (385 and 354), (383 and 333) respectively). Thus, under scoring systems 3(a) and (3b), the characteristics of 'moderately long' matches are not hugely different for matches between weaker servers and for matches between strong servers. It is interesting to note that the P(A wins) values for the matches between weaker servers under these scoring systems are much the same as their values under scoring systems 1(a) and 1(b), whereas, for matches between stronger servers, the P(A wins) values are less than those under systems 1(a) and 1(b) (as a result of the decreased means under systems 3(a) and 3(b)). It can be seen that system 3(b) is a better scoring system than system 3(a) for the stronger servers. However, for matches between weaker servers, although the value of P(A wins) for system 3(b) (i.e. 0.721) is less than its value under system 3(a) (i.e. 0.741), its value is nevertheless greater than the associated value for the stronger servers (i.e. 0.712). The other characteristics of system 3(b) (except efficiency) for matches between weaker servers are 'better' than those of system 3(a). Thus, from a statistical point of view, system 3(b) could reasonably be preferred to system 3(a).

Fourthly, considering scoring systems 4(a) and 4(b), the means, the standard deviations and the 98% points of scoring systems 4(a) and 4(b) are, as expected, less than those of systems 3(a) and 3(b). For the matches between stronger servers, scoring system 4(b) is better than scoring system 4(a) (whilst, for the weaker servers, it 'goes somewhat close' to being better). It is also noted that, for matches between weaker servers, although the value of P(A wins) for 4(b) (i.e. 0.706) is less than its value for system 4(a) (i.e. 0.722), its value is nevertheless greater than the associated value for the matches between strong servers (i.e. 0.702). Thus, it can be argued that system 4(b) could reasonably be preferred to system 4(a). It can also reasonably be argued that, if the fifth set is simply a match tie-break, it would seem appropriate to use no-ad games rather than advantage games during the first four sets.

Scoring systems 1(a) and 1(b) are not considered any further in this paper as they have large values for the mean, the standard deviation and the 98% point, particularly for the strong servers, and also because they are no longer in use.

For matches between strong servers, the 98% point for system 3(a) (385) is substantially less than the 98% point for system 2(a) (582), whilst the value for P(A wins) is a little less for system 3(a) (0.708 compared to 0.723). For matches between weaker servers, the decrease in P(A wins) under system 3(a) rather than under system 2(a) is miniscule (0.741 down from 0.743). It follows that system 3(a) is a very reasonable alternative to system 2(a).

We have noted above that, from a statistical point of view, system 3(b) might reasonably be preferred to system 3(a). On the other hand, system 3(b) involves the use of no-ad games, which some players, spectators and viewers might find less attractive than advantage games. Nevertheless, if a lower mean and a lower 98% point than those of system 3(a) were required, system 3(b) could reasonably be used.

It can be seen from table 1 that scoring system 4(a) has no overall advantage over system 3(b) either for matches between strong servers or for matches between weaker servers. Also, we have noted above that system 4(b) could reasonably be preferred to system 4(a). Thus, if a smaller average match duration is required to that under system 3(b), system 4(b) could be considered. Judgements about the appropriate average length (and variation) required for a particular tennis match or tournament, however, are best made by the tennis enthusiasts, and not by mathematicians.

Summarizing, the conclusions about the best of 5 sets scoring systems presently available, are

1. Systems 2(a) and 2(b) are only really applicable to the finals of important doubles events, where the winners are not required to play another match within the next few days, and in situations in which there is no match scheduled to follow.
2. Across all matches with both strong and weaker servers, the characteristics of 'moderately long' matches are not hugely different for the two scoring systems 3(a) and 3(b).
3. From a statistical point of view, all things considered, system 3(b) could reasonably be preferred to system 3(a), and system 4(b) could reasonably be preferred to system 4(a).
4. If the fifth set is simply a match tie-break, it would be appropriate to use no-ad games rather than advantage games throughout the first four sets.
5. If a lower mean and a lower 98% point than those of system 3(a) are required, system 3(b) could reasonably be used.
6. For a shorter match than under system 3(b), system 4(b) (rather than 4(a)) could be used.

A comparison with the best of 5 sets scoring systems using '50-40' games

The four best of 5 sets systems using '50-40' games (see columns (c) in table 1) are now considered, and the following observations made.

1. For matches between strong servers, scoring system 2(c) is a better scoring system than both 2(a) and 2(b), scoring system 3(c) is a better scoring system than both 3(a) and 3(b), and scoring system 4(c) is a better scoring system than both 4(a) and 4(b), and, perhaps surprisingly given the number of points played, scoring system 4(c) is even a better scoring system than both 3(a) and 3(b).

(i) **(ii)** **(iii)** **(iv)** **(v)**	**P(A wins)** **Mean** **Stand Dev** **Efficiency** **98% Point**	**(a)** $\mathbf{p_a = 0.77;}$ $\mathbf{p_b = 0.73}$	**(b)** $\mathbf{p_a = 0.77;}$ $\mathbf{p_b = 0.73}$	**(c)** $\mathbf{p_a = 0.77;}$ $\mathbf{p_b = 0.73}$	**(a)** $\mathbf{p_a = 0.62;}$ $\mathbf{p_b = 0.58}$	**(b)** $\mathbf{p_a = 0.62;}$ $\mathbf{p_b = 0.58}$	**(c)** $\mathbf{p_a = 0.62;}$ $\mathbf{p_b = 0.58}$
1	(i)	0.778	0.759	0.741	0.752	0.728	0.718
	(ii)	484.3	366.7	230.4	274.2	239.3	200.1
	(iii)	218.2	149.6	68.2	73.8	61.6	49.4
	(iv)	0.34	0.38	0.52	0.61	0.57	0.61
	(v)	1049*	755*	395*	442*	377*	306*
2	(i)	0.723	0.721	0.730	0.743	0.723	0.715
	(ii)	290.3	259.0	211.7	262.1	230.7	196.6
	(iii)	99.5	76.7	52.0	63.6	54.2	46.3
	(iv)	0.34	0.38	0.51	0.59	0.55	0.61
	(v)	582*	467*	323	395	341	290
3	(i)	0.708	0.712	0.727	0.741	0.721	0.715
	(ii)	272.0	248.9	210.0	261.0	229.9	196.3
	(iii)	60.7	55.7	48.6	61.6	52.7	45.7
	(iv)	0.32	0.36	0.50	0.58	0.55	0.60
	(v)	385	354	306	383	333	286
4	(i)	0.699	0.702	0.714	0.722	0.706	0.700
	(ii)	255.4	234.2	198.7	245.5	216.7	185.8
	(iii)	44.8	41.6	37.7	46.9	40.0	35.5
	(iv)	0.31	0.35	0.46	0.52	0.50	0.55
	(v)	333	308	270	335	291	253

Table 1: The five statistics for the twelve best of 5 sets scoring systems (* are simulated values)

2. Across all matches with both strong and weaker servers, scoring system 4(c) goes very close to being a better scoring system than system 4(b), and scoring system 3(c) goes very close to being a better scoring system than system 3(b).

3. The characteristics of system 3(c) are reasonably similar for the matches between strong servers and those between weaker servers, as are the characteristics of system 4(c). Thus, any match under system 3(c) is likely to have quite similar duration characteristics regardless of the serving strength of the players, and likewise, any match under system 4(c) is also likely to have quite similar duration characteristics.

4. Overall, systems 2, 3 and 4 using '50-40' games appear to be at least as good as these three systems when using no-ad games. Thus, the '50-40' game provides a real practical alternative to the no-ad game, when used for best of five sets matches.

A comparison of best of 3 sets systems 5, 6, 7 and 8, using deuce, no-ad and '50-40' games

The corresponding results for the best of 3 sets scoring systems are given in table 2, and the following observations are made.

1. Many of the comments and conclusions on the best of 5 sets systems apply to the corresponding best of 3 sets systems.

2. For matches between strong servers, system 6(b) goes very close to being a better system than 6(a), whilst scoring system 6(a) has a large standard deviation (93.1), and a large 98% point (480). Thus, it could be argued that system 6(a) is only really applicable to the finals of important doubles events, where the winners are not required to play another match within the next few days, and in situations in which there is no following match. It is noted that the corresponding 98% point for system 6(b) is 366, which is a much better value (than 480) for tournament play.

3. For matches between strong servers, system 6(c) is better than 6(a) and 6(b), system 7(c) is better than 7(b) (which in turn is better than 7(a)), and system 8(c) is better than 8(b) (which in turn is better than 8(a)). System 8(c) even goes close to being better than system 7(a).

4. As with the best of 5 sets systems, the efficiency of all of the (a) and (b) systems is low for matches between strong servers, whereas it is around 0.5 for the (c) systems. The best of 3 sets systems are, as expected, slightly more efficient than the corresponding best of 5 sets systems.

5. With the exception of efficiency (and the value of P(A wins)), the characteristics of system 7(a) are quite similar for matches between strong and weaker servers, so the length of matches under system 7(a) are not particularly dependent on the strength of the players' serves.

6. If it was felt that matches under system 7(c) and/or 8(c) were 'too short', then playing first to 7 games sets would increase the duration whilst maintaining efficiency and increasing the value of P(A wins). (It is noted that first to 6 games sets can be unfair (Pollard, 2005), and that the associated unfairness can be removed by playing first to 7 games sets.)

	(i) P(A wins) (ii) Mean (iii) Stand Dev (iv) Efficiency (v) 98% Point	(a) $p_a = 0.77$; $p_b = 0.73$	(b) $p_a = 0.77$; $p_b = 0.73$	(c) $p_a = 0.77$; $p_b = 0.73$	(a) $p_a = 0.62$; $p_b = 0.58$	(b) $p_a = 0.62$; $p_b = 0.58$	(c) $p_a = 0.62$; $p_b = 0.58$
5	(i)	0.730	0.713	0.698	0.707	0.687	0.678
	(ii)	298.2	225.3	141.2	168.3	146.5	122.4
	(iii)	165.0	112.1	48.9	51.6	42.8	33.9
	(iv)	0.36	0.40	0.55	0.65	0.60	0.65
	(v)	750*	533*	269*	295*	250*	200*
6	(i)	0.690	0.686	0.690	0.701	0.683	0.676
	(ii)	192.1	166.4	131.0	161.6	141.8	120.5
	(iii)	93.1	66.8	37.9	44.6	37.6	31.7
	(iv)	0.37	0.41	0.54	0.63	0.59	0.65
	(v)	480*	366*	225*	261	226*	188
7	(i)	0.669	0.672	0.686	0.697	0.681	0.675
	(ii)	166.3	155.2	128.6	160.0	140.7	120.1
	(iii)	40.3	37.0	32.6	41.4	35.3	30.7
	(iv)	0.34	0.38	0.53	0.62	0.58	0.64
	(v)	243	224	196	246	212	183
8	(i)	0.656	0.658	0.667	0.670	0.658	0.655
	(ii)	142.8	131.5	112.5	137.8	122.0	105.2
	(iii)	21.8	20.5	19.9	24.5	20.5	18.8
	(iv)	0.33	0.37	0.48	0.52	0.51	0.56
	(v)	187	174	158*	191	166	146

Table 2: The five statistics for the twelve best of 3 sets scoring systems (* simulated values)

CONCLUSIONS

This paper provides answers for those persons wanting to do something about the undefined length of tennis matches, where the scoring systems used lead to matches of unpredictable length, with considerable variation and exposure to excessive length, all of which affect the scheduling, television coverage and players' health. Experimentation with modified scoring systems is taking place in professional tennis, especially in men's doubles where eight different scoring systems have been used to date. All of these eight systems have been analysed in this paper, using the usual advantage scoring in each game, no-ad scoring, and the more efficient '50-40' game scoring system. Thus, twenty four scoring systems in total have been analysed and discussed.

Five statistics are sufficient to describe and compare tennis scoring systems. By using these statistics, one can gain an insight into the effects of selecting one scoring system for a tournament rather than another.

For best of 5 tie-break sets matches, the characteristics of ‘long’ matches are not hugely affected by whether advantage or no-ad games are used. If tournament organizers require fewer points in such ‘long’ (best of 5 tie-break sets) matches, a match tie-break fifth set and no-ad games throughout, might be used instead of five tie-break sets. Indeed, the results in this paper indicate that if a match tie-break is to be used for the fifth set, no-ad games are better (than advantage games) to use throughout the first four sets. For all of the best of 5 sets systems presently in use, the results show that the use of the ‘50-40’ game would be at least as good as using the no-ad game, making the 50-40 game a practical alternative to the no-ad game. For the best of 3 sets matches, similar comparisons exist between the various systems.

At the Grand Slam level, Wimbledon retains best of 5 sets for men’s doubles using scoring system 2(a) (4 tie-break sets and fifth set advantage). The other Grand Slams have all moved progressively to system 7(a) (best of 3 tie-break sets) in order to achieve a considerable reduction in the average length of matches, the standard deviation of the length, and the 98% point for the length of matches. Whilst Wimbledon may not particularly wish to achieve savings in average length and variation in length for doubles matches, the analysis in this paper shows that Wimbledon could achieve similar reductions to those in the other Grand Slam events but retain best of 5 sets structure by using system 4(c) (i.e. 4 tie-break sets and the fifth set a match tie-break, using ‘50-40’ scoring in the first 4 sets).

The ATP and WTA have sought even greater reductions in length and variation in length of matches, and are currently using system 8(b) (2 tie-break sets and third set a match tie-break, with no-ad scoring). If they used system 8(c) (same as 8(b) but using ‘50-40’ scoring), they would achieve further reductions in mean length, standard deviation, and in the 98% point, while efficiency would actually improve. Alternatively, they could use system 6(c) (which uses 50-40 games) and thus retain the 3 tie-break sets structure, and avoid the match tie-break.

In general, tournaments would like much greater control over the length of doubles matches, but different tournaments may want different average lengths of matches. This paper presents 24 scoring systems options that tournaments might use to increase the certainty that matches approximate that desired length. It is argued that it is not appropriate in this paper to be prescriptive about which scoring system to use in one tournament or another. However, tournament organizers who are interested in two or more scoring systems for possible use in a tournament can refer to the results in the body of this paper as a guide to making that decision.

References

Brown, A., Barnett, T. and Pollard, G. (2008) A recursion method for evaluating the moments of a nested scoring system. In J. Hammond and N. de Mestre (Eds.) *Proceedings of the Ninth Australasian Conference on Mathematics and Computers in Sport.* Mathsport (ANZIAM), Coolangatta, Queensland, Australia.

Miles, R. (1984) Symmetric sequential analysis: the efficiencies of sports scoring systems (with particular reference to those of tennis). *Journal of the Royal Statistical Society, B*. **46(1)**, p. 93-108.

Pollard, G. (1983) An analysis of classical and tiebreaker tennis. *Australian Journal of Statistics*. **25(3)**, p. 496-505.

Pollard, G. (2005) On solving an aspect of unfairness in the tennis doubles scoring system. *Chance*. **18(4)**, p. 17-19.

Pollard, G. H, Barnett, T., Brown, A. and Pollard, G. N. (2007). Some alternative men’s doubles scoring systems. In S. Miller and J. Capel-Davies (eds.) Tennis Science and Technology 3, International Tennis Federation Licensing (UK) Ltd, Roehampton, London SW15 5XZ. p. 301-309.

Pollard, G. and Noble, K. (2004) The benefits of a new game scoring system in tennis: the 50-40 game. In R. Morton and S. Ganesalingam (Eds.) *Proceedings of the Seventh Australasian Conference on Mathematics and Computers in Sport*. Massey University, Massey, New Zealand, p. 262-265.