ABSTRACT

A study was conducted on the tournament structure of women’s beach volleyball. The official world rankings were used in a logistic regression model to estimate the chance that each team has of winning a match. Match results were gathered from the 1998 and 1999 season, and testing showed that the official rating system provided a good estimate of performance for teams. The probabilities produced by the regression were then used as input to a simulation, which can be applied at any stage of a tournament to ascertain the probabilities of each team winning, or placing in any position. The simulation was used to investigate the efficacy of different tournament structures. The simulation was then used during the Olympic tournament, and the chance of each team winning was updated on a web page as the tournament progressed.

Keywords  Sports, Simulation, Logistic regression, Volleyball, Olympics

1. INTRODUCTION

Beach Volleyball is a recent addition to Olympic competition, first making an appearance in 1996 at the Atlanta games. The typical match is played between two teams, each with two players. The objective is to be the first team to score 15 and to be at least two points ahead. The most common style of tournament is double elimination, in which a team has to lose twice to be eliminated from the tournament. However the format used in the Sydney Olympic Games was changed to one that combined both single elimination and repeaches. Both tournament structures are seeded, so that highly rated teams do not play each other early in the tournament. The seedings are based on the ratings prior to the tournament and are maintained by the governing body of volleyball, the Federation Internationale de Volleyball (FIVB). Each tournament the player receives a certain number of rating points depending on the finishing position of their team. The “team” ratings used to rank the teams prior to each tournament is the sum of the two player’s individual ratings over the previous 12 months. The number of ratings points on offer for the winners of each tournament is dependent on the amount of prize money on offer. Teams finishing behind the winner get a proportion of the winner’s points depending on finishing position. Second receives 90% of the points allocated to the winner, through to the eight teams that finish in equal seventeenth place who receive 20% of the points assigned for first place. Teams that finish in 25th placing or lower receive no points.

To estimate the chances of winning in Australian rules football, Bailey [1] uses many predictors such as demographic data (age, weight, experience), performance statistics (kicks, marks, disposals, turnovers) and other data such as bookmakers’ prices. There are many variables that might be used as predictors of success in beach volleyball. In addition to those above, home advantage is known to be an important factor in international sport, particularly in the Olympics [2]. Many other factors, such as environmental conditions and the schedule itself could be added to the list. However time constraints and data available limit any study. This study was done as part of an undergraduate project for the subject Modelling in Sport, and was necessarily limited in scope. The ultimate aim was to provide content to the University web site giving updated daily predictions for the Sydney Olympic tournament, so the approaching Olympics also provided severe time constraints. In tennis [3] and soccer [4], Clarke & Dyte used official ratings to produce probabilities of match outcomes, which were then used to simulate a tournament. Since we could

1 School of Mathematical Sciences, Swinburne University of Technology, Melbourne (Email: sclarke@swin.edu.au). Submitted: February 26, 2001 and accepted: May 9, 2001 after one revision.
not obtain any demographic or performance data, we were attracted to the simplicity of this technique. It should also be noted that FIBV also ignore all other possible predictors, and only use the team ratings to determine seedings. As an undergraduate project this does not pretend to be at the cutting edge of research. However it does provide an example of successfully applying a previously documented method to a different area, and implementing the results. In this paper, we first use logistic regression to estimate the chance of each team winning a match, depending on their FIVB rating. These probabilities are then used to simulate past tournaments, and investigate the efficacy of different tournament structures. This is in contrast to previously reported simulations of tournament structures, where mathematically generated probability models are used. We then describe how the method was applied to the Sydney Olympic tournament.

2. DATA COLLECTION

To fit the model a considerable number of match results was required. The data consisted of ratings for the teams prior to each tournament and the results for each match, and was gathered from the FIVB website www.fivb.org. Data was collected from 11 World Tour tournaments in the 1998 and 1999 seasons, a total of 654 matches. A quarter of the data was kept for a holdout sample. The earliest 491 matches were used to fit the model, and the remaining 163 matches were used to test the predictive power of the model.

3. MODELLING INDIVIDUAL MATCHES

To see whether there is a relationship between the difference-in-rating between teams and their performance in a match a logistic regression was performed on the 491 matches. The general logistic equation is given by:

\[
\text{Chance of Team A winning match} = \frac{e^{\alpha + \beta\text{(Difference in rating)}}}{1 + e^{\alpha + \beta\text{(Difference in rating)}}} 
\]

(1)

Maximum likelihood estimates of the parameters \(\alpha\) and \(\beta\) can be determined using a statistical package. If the difference in rating is zero, we expect the model to give each team an equal chance of winning. This implies the parameter \(\alpha\) is equal to zero, giving the refined model:

\[
\text{Chance of Team A winning match} = \frac{e^{\beta\text{(Difference in rating)}}}{1 + e^{\beta\text{(Difference in rating)}}} 
\]

(2)

The logistic procedure of SAS/STAT version 8 applied to 491 observations (matches) showed the rating difference to be a highly significant (\(p<0.001\)) predictor of who would win the match, with \(\beta = 0.00119\) (stderr = 0.00014).

A graph of the logistic function is given in Figure 1. It shows how the chance of winning varies...
with the difference in ratings. The largest rating difference in the data collected was approximately 2000 points. As the rating difference increases, the chance of the better team winning approaches 1 asymptotically.

Figure 1 – Chance of Winning versus Rating Difference

The predictive power of the model can be tested on the holdout data set in various ways. The simplest is to count what percentage of matches was won by the favourite. 69% of the higher rated teams won their matches. One problem with this method is that it gives no weighting to the strength of prediction. When the model predicts that a team has a 51% chance of winning, it is treated the same as when the model predicts a team has a 99% chance of winning. It seems natural that the 99% prediction should be penalised more if it turns out to be incorrect. Dowell[5] gives an outline of a probabilistic football tipping competition where accuracy of predictions is rewarded by using a log-likelihood function. Each tipster is required to give an estimate of the probabilities of each team winning, and the score is the sum of the logarithm of the probabilities of the actual winning team. Maximum points are given to the tipster who gives a team a 100% chance of winning and that side actually wins; unfortunately if this tipster gets the result wrong his penalty will be infinitely negative points. A naive tipster who gives each team a 50% chance of winning is not affected by who wins; they will receive the same amount of points irrespective of the result. The tipster who most closely matches the true likelihood of the final result with his estimates should win this competition over a season. Forrest[6] also uses log-likelihood functions to compare football tipsters in English soccer.

The log-likelihood of the model’s tips for the holdout set is –92.8. To see how much information is contained in our model we can compare it with an uninformed, or naive, tipster who gives each team a 50% chance of winning a match. The log-likelihood for the naive tipster is -113.0. Bolton and Chapman[7] use a pseudo-R-squared to assess a multinomial logit model applied to horse racing. This is analogous to the multiple correlation coefficient in linear statistical models, and is given by:

\[ R^2 = 1 - \frac{L(\text{estimated})}{L(\text{naive})} \]

where \( L(x) \) is the log-likelihood function. The pseudo-R2 value gives the percentage gain in information of our model over the naïve tipster. For our model the pseudo-R-squared is 18%.

4. TOURNAMENT STRUCTURES

There are many ways of setting up the structure of a tournament. Round robins, knockouts, repechages, can all be used separately or in combination to produce a final ordering of teams. The structures we investigate here are those that are currently used in volleyball, being the double knockout and the Olympic tournament, and, as it is widely used in other sports such as tennis, the single knockout system.

In seeded single knockout tournaments, teams are seeded in some form and play each other in knockout fashion. The top two seeds are in opposite sections of the draw such that they cannot play each other until the final. To give the best teams the greatest chance of winning it is important that the seeding system is accurate[8]. In our simulations we have seeded every team in the competition, although in some sports only the top teams/players are seeded. A seeded knockout with 32 teams will take 31 matches to determine the winner. The seeded double knockout is the format used in international beach volleyball competitions. In the first round top seeded players will play the bottom-seeded players of the tournament, i.e.: P1vsP32, P2vsP31, …, P16vsP17.
Teams are allowed to compete until two losses are incurred. After one loss, the losing team moves to the losers' bracket and competes for two spots in the semi-finals with the two teams qualified from the winners' bracket. Once a team is in the losers' bracket they must play two matches per round to remain in the competition. From the Semi-Final stage onwards all matches are sudden death, with two teams from each bracket making up the semi-finals. The final match is played best of three sets. A double seeded knockout with 32 teams will take 61 matches to determine the winner. The first two rounds of a double knockout tournament are shown below.

Rd 1 – Start with 32 teams, 16 matches. This gives:
W – 16 teams move to Winners Bracket
L – 16 teams move to Losers Bracket

Rd 2 – Winners play Winners, Losers Play Losers
WW – 8 Winners Bracket
WL – 8 Losers Bracket
LW – 8 Losers Bracket
LL – 8 Eliminated

Rd 2.5 – 16 teams in the Losers Bracket play extra round
WLW or LWW – 8 Losers Bracket
WLL or LWL – 8 Eliminated

This leaves eight undefeated teams, and eight teams that have lost only one match. The 16 teams that have lost two matches have been eliminated. The tournament continues in the double elimination format until the semi-finals.

The structures that we have simulated are by no means an exhaustive list. World cup soccer uses a combination of a Round Robin with a Seeded Knockout. McGarry[9] found that it was not good for determining the best team but is reasonable for placing the no.1 team in the top three. In the McIntyre Final Eight system, used in some Australian rules football and rugby competitions, elimination after a loss depends on the results of other matches [10]. In the first round the two lowest seeded losers are eliminated, while the two highest seeded losers get a second chance.

5. SIMULATING TOURNAMENTS USING ESTIMATED MATCH PROBABILITIES

Previous studies on the properties of different tournament structures have usually not involved the exact structures that we are investigating. McGarry[9] uses simulated eight team tournaments to compare different structures. Edwards[11] investigates a different double elimination structure to the one used in beach volleyball. It was found that the seeded double knockout structure looked at in Edwards[11] is better than a seeded single knockout for producing a final finishing order that is closest to the order of the abilities of the teams. Both these studies use idealised probability models not related to actual teams. We wish to test the validity of these results with the data we have gained for a live tournament. The data needed for this to be done is the tournament structure and draw, and the individual match estimated probabilities that we compute with the logistic regression. This can be entered into a computer program that could play out a large number of simulated tournaments and work out a particular team’s chance of winning given their rating. This simulator can be used to simulate different tournament structures to see if there is any difference in the expected probabilities for each structure.

The double knockout structure was programmed into the computer and the ratings were used to give each team a chance of winning each match. The simulated winner of a match is determined by generating a random number and using that to determine the winner. The winner is then
passed through to the next round of the tournament, and the process is repeated until a winner is declared. By repeating this process thousands of times, the probability of each team winning the tournament can be determined. Collating the results of the simulated tournaments was done with a Visual Basic macro. The program can also provide estimates for the chance of finishing second, third, or fourth. The model can also be used to simulate the remainder of a tournament that is already in progress by entering the results of already completed matches.

The Osaka tournament from the 1998 season was then simulated. For the simulated tournament, results were collected on each team’s chance of winning or placing. Another simulation was carried out using a single knockout structure with the same teams. The purpose of this was to see if there is any difference in the expected outcomes for different tournament structures. Table 1 shows the results for simulations of the 1998 Osaka tournament.

| Seed | 1st place DE | 1st place SE | 2nd place DE | 2nd place SE | 3rd place DE | 3rd place SE | 4th place DE | 4th place SE | Total top 4
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>51.3</td>
<td>51.2</td>
<td>19.7</td>
<td>20.4</td>
<td>16.0</td>
<td>9.6</td>
<td>3.9</td>
<td>2.2</td>
<td>90.9</td>
</tr>
<tr>
<td>2</td>
<td>24.3</td>
<td>23.1</td>
<td>25.5</td>
<td>24.5</td>
<td>19.3</td>
<td>15.3</td>
<td>7.0</td>
<td>5.5</td>
<td>76.1</td>
</tr>
<tr>
<td>3</td>
<td>12.9</td>
<td>11.6</td>
<td>18.4</td>
<td>18.1</td>
<td>21.4</td>
<td>20.1</td>
<td>13.4</td>
<td>12.2</td>
<td>66.1</td>
</tr>
<tr>
<td>4</td>
<td>5.7</td>
<td>5.8</td>
<td>11.6</td>
<td>7.6</td>
<td>11.6</td>
<td>16.0</td>
<td>11.4</td>
<td>14.8</td>
<td>40.3</td>
</tr>
<tr>
<td>Others</td>
<td>5.8</td>
<td>8.3</td>
<td>24.8</td>
<td>29.4</td>
<td>31.7</td>
<td>39.0</td>
<td>64.3</td>
<td>65.3</td>
<td>-</td>
</tr>
</tbody>
</table>

In the double elimination tournament the top seed finished in the top four 91% of the time, whereas in the single elimination the top seed finished in the top four only 83% of the time. This confirms McGarry and Schutz [9] that the double elimination structure is advantageous for the top seeds. The simulation estimates that the top seed finishes first 51% of the time. Empirically, from 1996 to 2000 the top seed won 23 out 55 tournaments (42%). While not many trials, it is consistent with the simulations, as the difference is not significant.

6. THE SYDNEY OLYMPIC GAMES

The Sydney Olympic Tournament has a unique structure. Due to restrictions associated with the Olympics, only 24 teams were allowed to compete. Also, only two teams from each country were allowed to compete. This meant that some strong teams from USA and Brazil missed out on competing in the Olympics. The 24 teams initially play a single round. The losers of this first round play a knockout competition amongst themselves to progress to the next round with the first round winners. The knockout that the first round losers play consists of two rounds and produces three teams that rejoin the main draw. One extra “lucky loser” is also promoted to the main draw to bring the total teams left to 16. From here on, it becomes a single knockout tournament. Single elimination tournaments can only be run with 2^n teams (2, 4, 8, 16, 32, 64, 128… 2^n). For the Olympic tournament we updated the simulator daily to provide revised probability estimates as it progressed and posted them on the Swinburne Sports Statistics web page at [http://www.swin.edu.au/sport](http://www.swin.edu.au/sport). One interesting aspect of the Olympics was that in the initial round, 11 of the 12 higher ranked sides won their matches. This dominance of the higher ranked teams continues through to the quarterfinals where seven of the eight most fancied sides remained in the competition. There appears to be quite a disparity in the ability of the best and worst teams. From the structure of the Olympic tournament, the lowly ranked teams are assured of at least two matches in the tournament. This may be one object of the tournament design to encourage participation of the weaker countries.
It is interesting to note that the Australian team of Cook and Pottharst are estimated to have equal chance of winning the tournament with the Brazilians, Pires and Samuel, even though the Brazilian pair outrate the Australians by 200 points. The reason for this is that the Australians were given the number one seeding by the tournament organisers, which usually ensures an easier draw. Tournament organisers generally give home teams higher seedings for promotional purposes. The eventual winners, Cook and Pottharst from Australia were only the fifth ranked side going into the start of the competition and beat the highest ranked Brazilian pair in the final. In fact, modelling purely from the rankings gave the Australians only a 15% chance of winning that final match. This may show some inadequacies in our model where it comes to home ground advantage. Perhaps the Australians had a greater chance of winning than we anticipated due to the large partisan crowd and familiar conditions.

Home advantage has been the focus of many studies in previous years covering many different sports. Schwartz & Basky[12] found that home advantage exists in baseball, pro football, ice hockey and college basketball. The percentage of matches won by the home team varies from 53% in baseball to 64% in ice hockey and college basketball. We can provide a modest home advantage by giving the Australian teams an extra 250 rating points. Substitution in (2) shows this gives them a 57% chance of beating an otherwise equal opponent. Table 3 shows new estimates taking into account this small home ground advantage. The estimated chance of the Australians winning the tournament increases from eight to 11% when home advantage is taken into account. Again, the advantage of being given the top seeding is shown as they have a greater chance of winning the tournament than the team of McPeak-May, who have a higher rating.

---

Table 2 – Percentage chance of winning the gold medal at various stages of the competition.

<table>
<thead>
<tr>
<th>Team</th>
<th>Rating</th>
<th>Country</th>
<th>Initial chance</th>
<th>% Before Finals</th>
<th>% Before Semi-finals</th>
<th>% Before Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behar-Shelda</td>
<td>3180</td>
<td>BRA</td>
<td>55</td>
<td>56</td>
<td>74</td>
<td>85</td>
</tr>
<tr>
<td>Jordan-Davis</td>
<td>2676</td>
<td>USA</td>
<td>16</td>
<td>14</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>McPeak-May</td>
<td>2606</td>
<td>USA</td>
<td>10</td>
<td>12</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pires- Samuel</td>
<td>2430</td>
<td>BRA</td>
<td>8</td>
<td>8</td>
<td>15</td>
<td>-</td>
</tr>
<tr>
<td>Pottharst-Cook</td>
<td>2230</td>
<td>AUS</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>Takahashi-Teru Saiki</td>
<td>1952</td>
<td>JPN</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>Others</td>
<td>2</td>
<td></td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3 – Percentage chance of winning the gold medal at various stages of the competition with home ground advantage.

<table>
<thead>
<tr>
<th>Team</th>
<th>Rating</th>
<th>Country</th>
<th>Initial chance</th>
<th>% Before Finals</th>
<th>% Before Semi-finals</th>
<th>% Before Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behar-Shelda</td>
<td>3180</td>
<td>BRA</td>
<td>53</td>
<td>54</td>
<td>71</td>
<td>77</td>
</tr>
<tr>
<td>Jordan-Davis</td>
<td>2676</td>
<td>USA</td>
<td>16</td>
<td>14</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>McPeak-May</td>
<td>2606</td>
<td>USA</td>
<td>10</td>
<td>12</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pires- Samuel</td>
<td>2430</td>
<td>BRA</td>
<td>8</td>
<td>7</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>Pottharst-Cook</td>
<td>2480</td>
<td>AUS</td>
<td>11</td>
<td>11</td>
<td>15</td>
<td>23</td>
</tr>
</tbody>
</table>
7. CONCLUSION

Presented here is a model used to predict the results of beach volleyball matches and tournaments. The model uses the official world ratings in a logistic regression to estimate a team’s chance of winning each match. This method is useful when time or data constraints preclude more complicated models. On a holdout set of matches the model performed better than a naive tipster. The output from the regression model can be used to generate probabilities for input to a computer simulation of entire tournaments.

Single and Double Elimination tournament structures were investigated using the simulator to compare their various characteristics. Single Elimination tournaments were found to have a greater amount of luck involved and give lower skilled teams a better chance of winning than they would have in a double elimination tournament.

The Sydney Olympics introduced a new tournament structure. This was simulated, with the model’s predictions posted to a web-site as the Olympics progressed. The winners of the tournament were Australian which highlighted one of the weaknesses of the model; it didn’t account for home advantage. A revised model was produced which accounted for the home advantage effect by boosting the ratings of local teams.

In some golf pairs tournaments a high and low handicapper will generally outscore two players with the same average handicap [13]. A simple extension of our model would be to include each player's ratings as predictors, rather than the total rating. This might lead to alternative ways of combining individual ratings to produce team ratings, or assist in selecting players to combine into good teams. The degree to which individual player ratings predict team success might be used as a measure of the teamwork involved in a sport. Thus one might expect progressively worse predictions as one fitted lawn bowls pairs (where each player bowls independently of their partner), table tennis doubles (where the rules prescribe that partners must alternate in hitting the ball), to tennis doubles (where players have the freedom to move around the court and decide who will return the ball). However, such a study might require more mixing of teams than might exist in elite sport, where many teams keep the same members for extended periods.

Although there are many other factors that can influence matches, and could be incorporated in a more sophisticated model, in beach volleyball the official ratings do provide a good estimate of the relative strength of teams. This is reassuring, as the ratings are used in the seeding of tournaments, which requires accurate assessments of team abilities.

REFERENCE


