

Research article

Retired matches among male professional tennis players

Kristijan Breznik ¹✉ and Vladimir Batagelj ²

¹ International School for Social and Business Studies, Celje, Slovenia; ² Faculty of Mathematics and Physics, University of Ljubljana, Slovenia

Abstract

The aim of this study was to explore the effect of characteristics of various games and players on the proportion of retired tennis matches in the Open Era of tennis. The data included over 420,000 matches played among 17,553 tennis players in the period from 1968 to the end of 2010. The influence of the surface type was clearly confirmed, with the proportion of retired matches being higher on hard and clay courts compared to grass and carpet surfaces. Similarly, more retired matches were observed in outdoor venues than in indoor ones. The impact of other variables, tournament types, rounds at which the game was played and both players' ranks, is more ambiguous. Our interpretation of the obtained results is presented in the paper. Network analytic methods were applied to extract players with the most retired matches in their careers. Eventually, we defined a group of top tennis players and gave a more precise insight into retired matches in that group. Correspondence analysis was used to visually display the two-mode network of top players and the proportion of retired matches by surface type.

Key words: Tennis surface, individual sport, network analysis.

Introduction

Tennis injuries are well analyzed at different playing levels (from recreational and collegiate levels to the professional level) and from various perspectives such as occurrence, aetiology and prevention (Pluim et al., 2006). Retired matches in professional tennis are commonly closely related to injuries. Specifically, in a glossary of tennis a retirement is defined as "a player's withdrawal during a match, usually due to injury, causing the player to forfeit their place in the tournament". Moreover, a tennis match is officially recorded as retired (see for example Chapter X of the 2011 ATP Rule Book) when a losing player retired due to illness or injury after the match began. In Pluim et al. (2007) medical services on the ATP Tour are described in more detail. Although the outcomes of retired matches might have a substantially different effect compared to the outcomes of other matches, in recent articles on tennis, retired matches were not analyzed separately nor were they excluded from the analysis (e.g. Del Coral, 2009; Del Coral and Prieto-Rodriguez, 2010; Koning, 2010). In the analysis of Pollard et al. (2006) matches in which one player retired before the match was finished were called "spurious data" and were correctly omitted.

The aim of our paper was to explore the influence of match and player characteristics on the quantity and proportion of retired matches among professional tennis

players. The effects of playing surface on the frequency of injuries in tennis matches have been studied abundantly (e.g. Bastholt, 2000; Girard and Millet, 2004; Fernandez-Fernandez et al., 2010; Nigg and Segesser, 1988). As reported in O'Donoghue and Ingram (2001) and in Fernandez et al. (2006) the court surface influences the match activity among the elite level tennis players. A hard surface is believed to be the most high-risk surface for injury incidence. Significantly fewer injuries are generally observed on clay courts or synthetic surfaces compared to hard courts (e.g. Girard et al., 2007, 2010; Murias et al., 2007). Consequently, the highest proportion of retired matches is expected on hard courts and the lowest proportion on clay courts.

Furthermore, other variables could influence the number of retired matches in tennis, for example the type of tournament and the round being played. We conjectured that the number of retired matches was inversely related to the importance of matches played. More precisely, in major tennis tournaments and similarly at a later stage of the tournament, fewer retired matches were expected. The reason could be that it is expected that players will persevere in their efforts to win longer, more important types of matches than participating in less important tournaments and matches.

Our next hypothesis refers to the strength of both opponents playing the game. The strength of a professional tennis player is measured in ATP ranking points. As a result, each professional tennis player is ranked on the ATP ranking list which is typically released every month (there are some exceptions, i.e. the weeks when Grand Slams are played, New Year's holidays, etc.). Research of Del Coral and Prieto-Rodriguez (2010) confirmed that the difference in rankings of individual players is the most explanatory variable influencing the outcomes of tennis matches. Furthermore, rank differences are more important as we move to the top of the distribution of players. Whereas higher ranked players are usually less likely to suffer defeat against lower ranked players, we assume that higher ranked players could occasionally find the cause of an impending defeat in injury problems and consequently they could end the match prematurely with a retirement. Considering this, we anticipated a higher number of lost retired matches by higher ranked players.

Tennis is a sport that is typically played outdoors and during most of the competitive parts of the year elite tennis players are exposed to warm or hot ambient conditions. An increased body temperature (hyperthermia) and less than optimal hydration (hypohydration) have been

shown to have an influence on the health of tennis players and limit their performance (Kovacs, 2006; Magal et al., 2003). Nonetheless, some tournaments are played indoors where conditions are usually more suitable for players. A lower proportion of retired matches was therefore expected indoors than outdoors.

Using network analytic methods, in the second part of the study professional tennis players with the highest number of prematurely ended matches were identified and classified. Eventually, our focus in that part of the paper was on the retired matches of the best players who have ever played the sport of tennis.

Methods

Database

The Open Era of tennis began in 1968 when the Grand Slam tournaments agreed to allow professional players to compete against amateurs. Therefore, the data used in the study included 420,489 tennis matches played among male professional tennis players from 1968 to the end of 2010 and were obtained from the website of the Association of Tennis Professionals - ATP (<http://www.atpworldtour.com/>).

A minority of matches, exactly 10,801 or 2.57% of all matches, finished due to the retirement of a player. Additional variables obtained directly from the game data set (match characteristics) are the following: *date* of the match (day, month and year), *result* of each match by sets, *surface* type (clay, hard, carpet or grass), *venue* of the match, the *round* in which the match was played and the *type* of tournament classified by the ATP. The data originally contained 15 tournament types which were merged into five clusters: Grand Slams, ATP World Tour tournaments, ATP Challenger Tour tournaments, Futures and Team tournaments. The latter include Davis Cup and World Team Championship matches. ATP World Tour tournaments also include, among other things, all three current Masters Series tournaments (1000, 500 and 250), Barclays ATP World Tour Finals (previously known as the Tennis Masters Cup tournament) and some already discontinued tournaments such as the Grand Slam Cup and Grand Prix tournaments. For each match we also have the information on whether it was played outdoors or indoors.

Altogether, 17,553 tennis players were included in the analysis. Their nationality as defined in Geyer, 2010 (subdivided into Eastern and Western Europe, North and South America, Africa, Australia and Asia) was included as a variable. Furthermore, available ATP rankings of individual players for all periods from the start of official rating calculations in August 1973 to the end of 2010 were obtained from the ATP website (there was some inconsistency in rating releases until 1983).

To obtain a more detailed overview of retired matches among the best players on the tennis circuit we have defined the *top players* as those players who have been number one in the ATP rankings for at least one week in their career or those who have won at least one Grand Slam tournament since 1968. Overall, 53 tennis players have been characterized as top players, seven of

whom were still active on the ATP Tour at the end of 2010, i.e. Juan Martin Del Potro, Novak Djoković, Roger Federer, Juan Carlos Ferrero, Lleyton Hewitt, Rafael Nadal and Andy Roddick.

Procedures

To assess the influence of explanatory variables on the proportion of retired matches the data were analyzed using descriptive and inferential procedures. Chi-square tests were performed to establish whether there were statistically significant differences in the distributions of retired and regular matches in relation to other variables. Statistical significance was set at 5%.

A more precise insight into retired matches at the individual level was obtained using network analytic methods. A network is, in general, defined with a set of actors (represented by *vertices* or nodes) and a relation (represented by *links*) between them (Wasserman and Faust, 1994). Links may either be directed (*arcs*) or undirected (*edges*) and weighted or unweighted. By using additional data about vertices and links, networks representing matches were provided as described below.

In the *network of (retired) matches* actors are individual tennis players in our data set. Two players are related (linked with an arc) if they played at least one (retired) match in the observed period. The winning player is represented by the initial vertex, while the losing player is represented by the terminal vertex of an arc. The resulting network is directed and weighted by the number of (retired) matches of the player at the initial vertex to the player at the terminal vertex. For additional purposes an undirected network of (retired) matches was created by merging arcs between a pair of players into an undirected link (an edge). Edges in the undirected network of (retired) matches were weighted by the total number of (retired) matches played between two opponents.

Similarly, a two-mode network, *top players x surface network*, consisting of two sets of entities (players and surface types) was obtained. The presentation of a two-mode network was accomplished using correspondence analysis (Clausen, 1998; Greenacre, 2007). This method is suitable for a graphic display of the two modes in the affiliation network (Carrington et al., 2005).

Programs for collecting data from the ATP website and producing networks were written in R (R Development Core Team, 2011) which was also used for the statistical analysis. For the analysis of networks we used the Pajek program (Batagelj and Mrvar, 1996-2011).

Results

The proportion of retired matches increased over the observed period as can be seen on the Figure 1. Until 1990 the proportion of these matches was under 1.5%, but since 2002 it has been constantly over 3%. The year 1990 was an important turning point in the history of tennis as the Grand Prix events were discontinued and the ATP Tour was born. However, this should not be the main reason for the increase in retired matches after that year. To eliminate the influence of different tournament structure, residuals of retired matches at Grand Slam

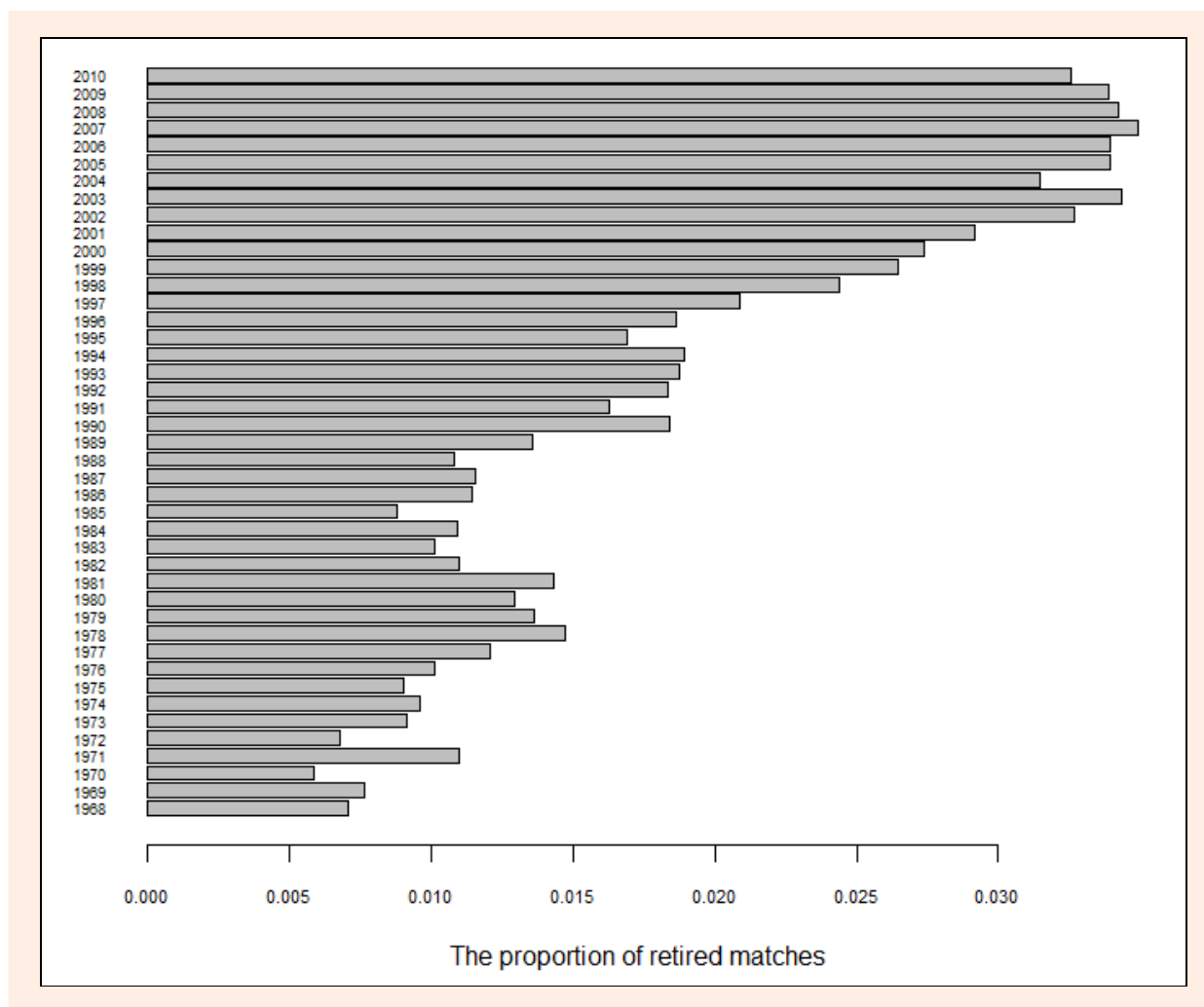


Figure 1. The proportion of retired matches by years.

tournaments were analyzed separately as the structure of Grand Slams seems to be the most stable of all tournament types and has remained virtually the same since the beginning of the Open Era of tennis in 1968. Figure 2 displays standardized residuals of a chi-square test of retired matches in Grand Slam tournaments.

Eight residuals before 1990 and seven residuals after that were statistically significant at the 0.05 level. The number of retired matches in these 15 years varied significantly from their expected frequency. Eight times before 1990 the observed frequency of retired matches was below the expected frequency and, conversely, after 1990 the observed frequency of retired matches was above the expected frequency. The results clearly confirm the year 1990 as the milestone in terms of retired matches also in Grand Slams.

Analysis per variable

Surface

Frequencies of all matches and retired matches separately (the latter also in percentage form) played on an individual type of surface are given in Table 1. Clay and hard courts were by far the most frequently used surfaces (46.98% of all matches were played on clay and 38.96% on hard courts which brings us to a total of 85.94% of all

matches). To adjust for this discrepancy the analysis of proportions by surface type is given. Nevertheless, the proportion of retired matches was significantly higher on clay and hard courts compared to the other two types of surface.

Table 1. Frequencies of all matches and retired matches by surface type.

Surface	All matches	Retired matches	Retired matches in %
Carpet	34,375	641	1.86
Clay	197,543	5,484	2.78
Grass	23,212	382	1.65
Hard	163,803	4,284	2.62
Unknown	1,556	10	0.64
Total	420,489	10,801	2.57

Furthermore, a 2 (type of match: regular, retired) x 4 (surface: carpet, clay, grass and hard) chi-square test was performed. The correlation between both variables was statistically significant ([chi-square] (3; n = 418934) = 181.882; p < 0.001). The number of retired matches was higher than the expected count on hard and clay courts and lower than the expected count on carpet and grass courts. Our hypothesis about court surface was therefore only partially confirmed. The assumption about hard

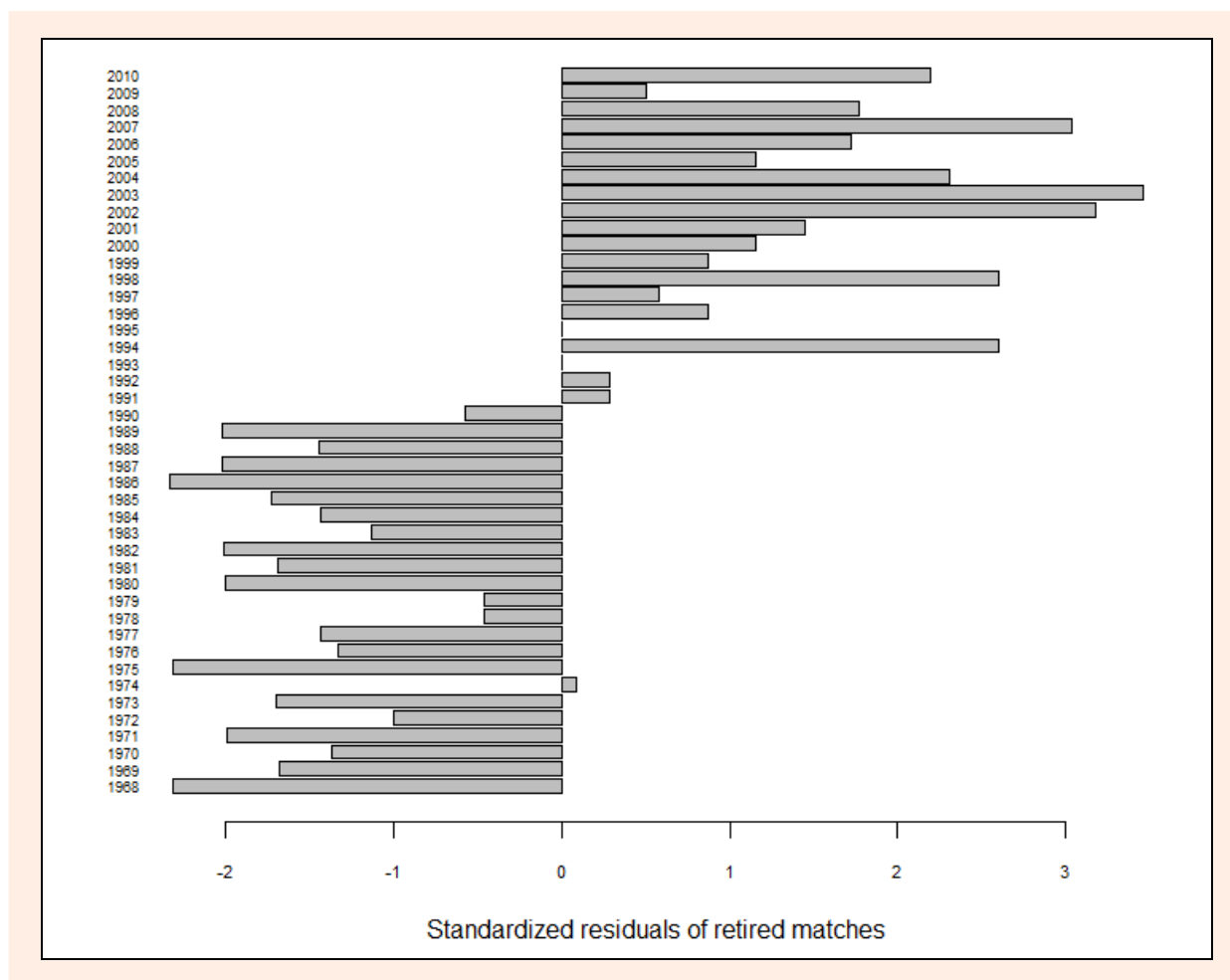


Figure 2. The standardized residuals of a χ^2 -test in Grand Slam tournaments by years.

Table 2. Frequencies of all matches and retired matches by tournament type

Tournament type	All matches	Retired matches	Retired matches in %
Grand Slams	22,037	533	2.42
ATP World Tour	125,241	2,014	1.61
ATP Challengers	110,825	2,886	2.60
Futures	144,418	5,187	3.59
Team	7,167	181	2.53
Total	420,489	10,801	2.57

Tournament types

The proportion of retired matches by tournament varied from 1.61% in ATP World Tour tournaments to 3.59% in Futures tournaments (see Table 2). Team type tournaments differed significantly from individual type tournaments and moreover, it was hard to determine the quality of team type tournaments. Therefore, our focus in this section was on individual types of tournament only. Grand Slams, as the most important tennis tournaments, had the second lowest proportion of retired matches. In recent years ATP World Tour tournaments (especially Masters 1000 series) had also been of a very high quality. Fewer players are able to compete in these types of tournament compared to Grand Slams where 128 players are in the initial draw. Matches played as part of Masters 1000 series were thus at least comparable to the matches played as part of Grand Slams. Another reason for the

higher number of retired matches in Grand Slam tournaments is that 'best of five' matches are played at those four majors only. As a result of this, matches at Grand Slams last longer than matches at other types of tournament and, consequently, more injuries can be expected. ATP Challengers and particularly Futures are less important tournaments and the proportion of retired matches was higher than in the above mentioned two more important types of tennis tournament. Our hypothesis about tournament types was thus confirmed. The proportion of retirements from matches was indeed inversely related to the importance of the tournament.

Rounds

Tournaments in tennis are organized, with very few exceptions, as single-elimination tournaments where the loser of each match is immediately eliminated from winning the championship or first prize in the event. However, the initial number of players (and consequently the number of rounds) in these types of tennis event varies from 28 to 128. The proportion of retired matches by round is displayed in Table 3. Since 2007 results of qualifying matches (labeled in Table 3 as Q1, Q2 and Q3) are included in the data set. Round robin (RR) tournaments are played in Barclays ATP World Tour Finals and in Team type tournaments.

A significantly lower proportion of retired matches

than expected can be observed in the round comprising the best 64 players and in the final. On the other hand, a significantly higher proportion of retired matches was found in the round comprising the best 32. Our hypothesis that the proportion of retired matches would be lower at the latter stages of the tournament was thus not supported.

Table 3. Frequencies of all matches and retired matches by round.

Round	All matches	Retired matches	Retired matches in %
Q1	9,984	271	2.71
Q2	6,310	172	2.73
Q3	2,831	62	2.19
R128	11,558	295	2.55
R64	29,800	546	1.83
R32	179,294	4,905	2.74
R16	91,092	2,229	2.45
Q	46,024	1,242	2.70
S	23,326	631	2.71
F	11,855	261	2.20
RR	8,415	10,187	2.22
Total	420,489	10,801	2.57

Rankings

On the basis of the difference in the rankings of both players, matches were clustered into five groups (see Table 4). 21.41% of all matches (90,012 matches exactly) were not in the resulting split as they had been played before ranking lists were created in 1973 or at least one of both players was not ranked when the match was played. Matches included in the ranking's analysis are predominantly more recent, which is the reason for a higher proportion of retired matches (2.82%) compared to the entire data set (2.57%).

Table 4. Frequencies of all matches and retired matches by difference in players' rankings and a quotient of retired matches by higher ranked players to retired matches by lower ranked players.

Rank difference	All matches	Retired matches	Retired matches in %	H _R /L _R
Under 25 (incl)	51,354	1,287	2.51	1.0158
26-60 (incl)	58,159	1,526	2.62	1.0000
61-120 (incl)	62,301	1,678	2.69	0.9265
121-250 (incl)	63,805	1,930	3.02	0.9674
Over 250	94,858	2,896	3.05	1.1201
Total	330,477	9,317	2.82	1.0169

The proportion of retired matches increases proportionally with a bigger difference in both players' rankings. A bigger difference in the rankings of both players resulted in a higher proportion of retired matches. In each group a quotient of retired matches by higher ranked players to retired matches by lower ranked players was calculated. Our hypothesis that higher ranked players retire more frequently than lower ranked players was supported only partially (in the first and last groups) and has been, in general, rejected.

Outdoor versus indoor venues

The majority of tennis matches in the data set (over 83% of all matches) were played outdoors. We assumed that unfavorable outdoors conditions resulted in a higher pro-

portion of retired matches. Indeed, the proportion of retired matches outdoors was 2.67% and the proportion of retired matches indoors was 2.07%. The hypothesis has been additionally confirmed using a Fischer exact test ($p < 0.001$).

Network analysis

Retired matches were analyzed at the individual level using the network of retired matches. 164 professional tennis players were involved in 20 or more retired matches in their careers. The highest number of retired matches is associated with a lesser known Mexican player Daniel Gorza. In this way he lost 29 and won 10 matches, both together representing almost 8.8% of the matches that he ever played. Among the top players the American player Jimmy Connors was involved in 29 retired matches (1.7% of all his matches). In fact, Connors also played the highest number of matches in our network (exactly 1,545 matches). Other top players involved in at least 20 retired matches in their career are: Andre Agassi (involved in 23 retired matches), Albert Costa (26), Juan Martin Del Potro (24), Novak Djoković (24), Juan Carlos Ferrero (32), Gaston Gaudio (20), Lleyton Hewitt (22), Goran Ivanišević (26), Ivan Lendl (21), Rafael Nadal (23), Marcelo Rios (24), Andy Roddick (22) and Marat Safin (20).

To determine the players who retired from the most matches among themselves or, in the words of network analysis, to detect cohesive subgroups in the network of retired matches, we applied a mixture of two techniques. Firstly, the valued core method also known as the *S-core method* (Batagelj and Zaveršnik, 2011) was used. A subset of vertices C determines the S-core at level t if and only if for each vertex in the set C the sum of weights of links of the vertex to other members of C is at least t . C also has to be the maximal set with this property. The highest valued core in the network of retired matches was the S-core at level 7. It was decided to analyze the S-core at level 4. An important feature of valued cores (and cores in general) is nesting – the S-core of the network of retired matches at level 4 also includes all S-cores of levels higher than 4. Specifically, the S-core at level 4 includes a group of all tennis players involved in at least four retired matches with other players in this group.

Whereas a lot of players (exactly 1,301) were included in the S-core at level 4, an additional (second) technique called the *islands approach* was used on the obtained S-core at level 4. The island approach detects subnetworks with stronger internal cohesion to its neighbors. A line island of size between k and K is a weakly connected subnetwork of the selected size in the interval $[k, K]$ where links joining vertices from the island to their neighbors outside the island have weights lower than the values of links of a spanning tree inside the island. The size of the islands in the S-core at level 4 was set between $k = 3$ and $K = 100$. Figure 3 displays all 26 obtained islands containing 120 tennis players. The players are colored by the regional partition (Geyer, 2010), red color being used for Eastern European players, grey for Western Europeans, green and blue for South and North American players, respectively, orange for Austra-

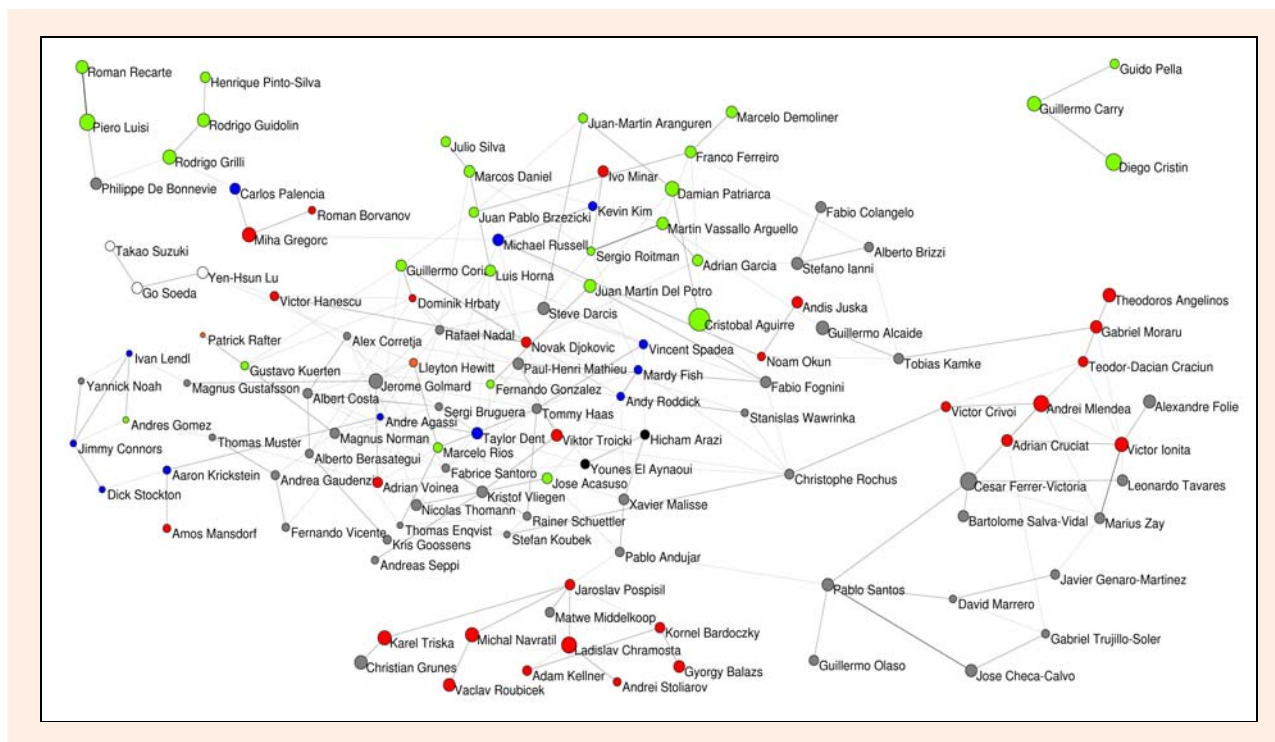


Figure 3. Islands in S-core at level 4.

lian players, black for African players and white for Asian players. The size of vertices representing players is proportional to the proportion of matches they retired from in their career. Places of vertices in Figure 3 were determined using the Kamada-Kawai algorithm (De Nooy et al., 2011). Only minor manual reallocations of some vertices were applied to obtain a clearer picture.

Figure 3 consists of two weak components – two isolated sets of players. A component in the top right part of the figure consists of only three lesser known South American players. However, the main component is more multinational. Most players in Figure 3 are from Europe and both Americas. There are only three Asian players and two Australian and two African players. The vertex representing an Argentinean player Cristobal Aguirre is the largest as he retired from 17.9% of his matches, followed by a Spaniard Cesar Ferrer-Victoria (10.8%) and another Argentinean Diego Cristin (10.0%).

In the central part we can identify some top players, i.e. Andre Agassi, Rafael Nadal, Novak Djoković, Marcelo Rios, Sergi Bruguera, Albert Costa and Lleyton Hewitt. Top players from the eighties and from the beginning of the nineties are placed in the left part: Ivan Lendl, Yannick Noah, Patrick Rafter, Gustavo Kuerten, Andres Gomez, Jimmy Connors and Thomas Muster. The only remaining top player in Figure 3, Juan Martin del Potro, is placed at the top of the figure. Some other parts of Figure 3 are also interesting. In the top left part we can identify another South American group with two Venezuelans, Recarte and Luisi, and three Brazilians, Pinto-Silva, Guidolin and Grilli. At the bottom a group of mostly Czech and Hungarian players can be found. An experienced Czech tennis player Jaroslav Pospisil connects them to the main part of the large component. Eventually, the right and bottom right parts are reserved for another European

group of players, mainly from Romania and the Iberian Peninsula. Three players, a Spaniard Pablo Santos, a Romanian Victor Crivoi and a German Tobias Kamke, represent a bridge to the other players.

Furthermore, a two-mode network of top players and the proportion of their retired matches by surface was presented using correspondence analysis. We selected surface type as the second mode since the analysis of explanatory variables indicated the surface as the most distinctive variable in relation to the number of retired matches. The first two dimensions together accounted for 73.63% of total inertia and are displayed in Figure 4 which uses principal coordinates for top players and standard coordinates for playing surfaces. Thus, in this figure each player is at the weighted centroid of the playing surface(s) on which he retired at least once during his career. Weights are determined as the proportions of the number of retired matches and the number of all matches played on the individual surface. Consequently, players are placed close to the surface type if these players had retired from a large proportion of matches on this surface. To improve readability in Figure 4 overlapping names of players were relocated slightly.

A majority of top players in Figure 4 come from Europe and both Americas, while other top players included six Australians and one African player. Top players are colored by the regional partition used already in Figure 3. Six top players, i.e. Arthur Ashe, Roger Federer, Vitas Gerulaitis, Rod Laver, Ken Rosewall and Mats Wilander, did not retire from even a single match during their career and are therefore not included in Figure 4. There are no Asian top players in the Open Era of tennis. Font sizes of players' names in Figure 4 are proportional to the proportion of matches they retired from in their career.

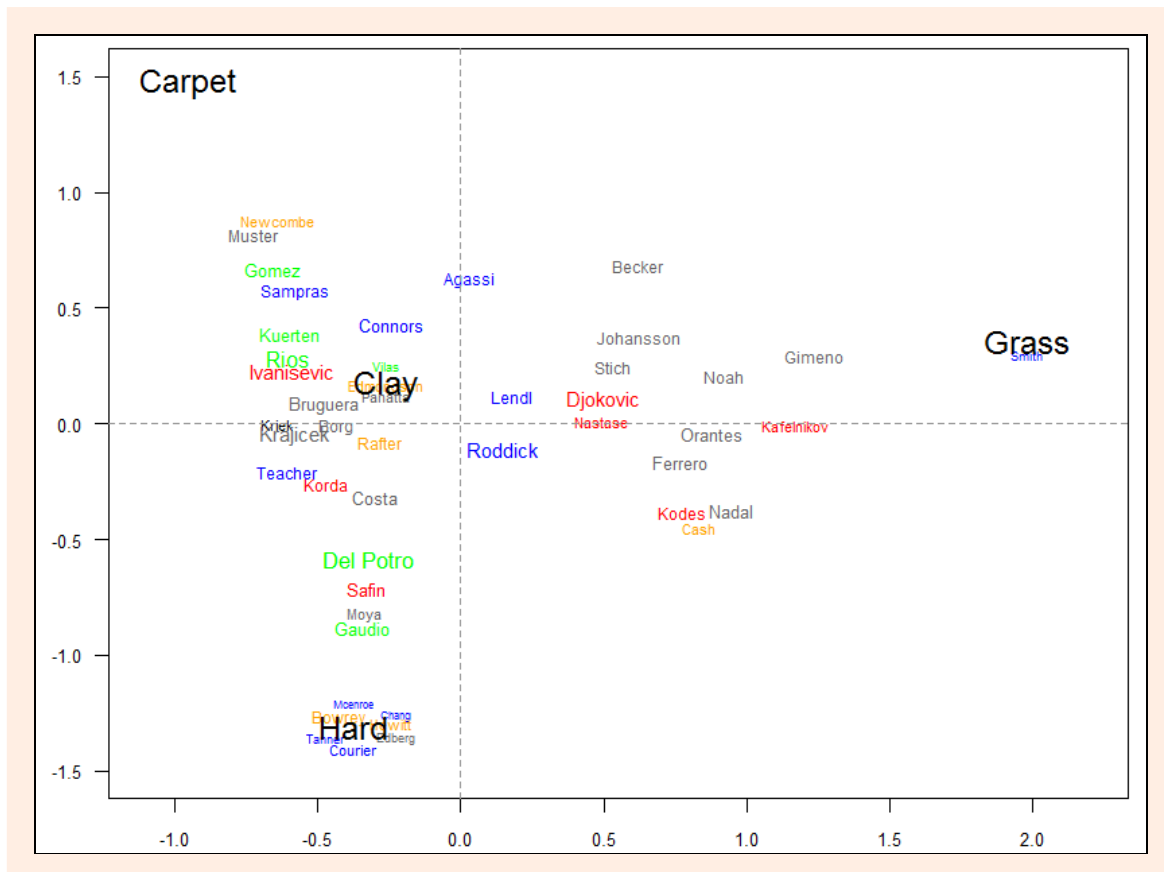


Figure 4. Two-mode network presentation of top players (principal coordinates) and playing surfaces (standard coordinates) using correspondence analysis.

Some interesting details can be seen in Figure 4. First dimension contrasts between European and North American top players, and the grass surface on the right side of the figure, and the other three surfaces with the rest of the players on the other side. Nevertheless, many Europeans and North Americans can be found on the left side, but there is just one Australian top player on the right. The clay surface is placed near the center of the figure and a majority of top players are gathered in the vicinity. The second largest group is gathered around the hard surface. On the contrary, only a few players are placed around the grass surface and none of them are around the carpet surface. A more detailed analysis has shown that seven top players retired from matches on hard courts only, three top players on clay courts only, one on grass courts only and none of the top players retired only on carpet. Observing active top players exclusively (i.e. Del Potro, Djoković, Roddick, Ferrero and Nadal), it could be concluded that they have retired from a relatively high number of matches as their names are written in large fonts. In this context, the achievement of Federer, the only active top tennis player not to have retired from a single match in his whole career, is particularly worthy of respect.

Discussion

The playing surface has been shown to influence injury rates dramatically in tennis. In general, a hard surface is believed to be the most high-risk surface for injury inci-

dence. However, on slow surfaces, such as clay courts, players play significantly longer rallies than on any other surface. For example, rallies on grass are significantly shorter (Barnett et al., 2006). Results of our paper represent strong evidence that playing tennis on hard and clay courts increases the incidence of retirements from matches on the professional tennis circuit. A high proportion of retired matches on hard courts was expected based on the results of previous studies. On the other hand, a high proportion of retired matches on clay courts could be a result of longer and more exhausting matches on the clay surface.

In papers published so far, no attention has been given to the influence of other variables included in our research – besides surface – on injury incidence in tennis. Our findings suggest that tournament quality is inversely related to the number of retired matches. The round at which the match is played does not seem to influence the latter significantly.

The establishment of a ranking system is an important achievement of the Open Era of tennis. Higher-ranked players win their tennis matches roughly at a rate of 71.5% (Del Coral and Prieto-Rodríguez, 2010). The authors obtained this result by analyzing Grand Slam tournaments from 2005 to 2008. The percentage is slightly higher than that of our data set (63.9%), but the discrepancy could be a result of our broader data set and a much longer period of analysis. In our study a rank difference was proportional to the proportion of retired matches – bigger rank differences resulted in a higher proportion

of retired matches. Interestingly, the quotient of retired matches by higher-ranked players to retired matches by lower-ranked players behaved in an unexpected manner indicating that the basic rank placement (above or below the opponent) does not have an effect on whether the match would result in retirement or not. However, if the rank difference was bigger than 250 places the chances that the higher-ranked player retired from the match were significantly higher (the quotient of retired matches by higher-ranked players to retired matches by lower-ranked players was approximately 1.1201).

Eventually, body effects as hyperthermia and hypohydration are more extreme when tennis matches are played outdoors in warm and hot environments. Therefore, we expected fewer retired matches among the matches played indoors which was confirmed in the analysis.

Conclusion

In the paper we studied the influence of various variables on the number/proportion of retired tennis matches. In compliance with previous studies on injury incidence and physical workload during a tennis match on different surfaces, hard and clay courts have been confirmed as the most risky surfaces for ending a match prematurely by a player's retirement. Common extreme outside conditions have been proven to be an additional reason for a higher number of retired matches in tennis. The results indicate that the difference in rankings of both players is proportional to the incidence of retired matches. On the other hand, tournament quality seems to be inversely related to the latter. It is therefore important not to consider a retired match only as a consequence of injury.

Methods used in the analysis have some direct consequences and implications. Currently, the US Open is the only Grand Slam tournament where a tie-breaker game at 6 games all in the fifth set is played while at the other three Grand Slams an advantage fifth set is played. We think that, in reference to the present analysis, on the clay courts of Roland Garros a tie-breaker game in the fifth set should also be considered. Recently, some retired matches have been associated with suspicious betting patterns (ESPN, 2007). The network analytic approach applied in the study could be used to identify and further investigate players involved in an excessive number of retired matches. Moreover, the approach is able to determine groups of players with stronger internal cohesion, in our case we were able to determine groups of players who retired from more matches to other players inside the group compared to the other players.

However, the study also has its limitations. It was conducted over a very long period of time which is why important tournaments and tournaments with a rich tradition (like Grand Slams and the Davis Cup) were more evenly represented compared to less important tournaments. Since in recent years all types of tournament have been covered, a more precise analysis of retired tennis matches will be possible in the future. Furthermore, only male tennis matches were analyzed in the paper. Future research could also include the women's tennis circuit and

matches played between doubles teams (including mixed doubles), and could test whether there are important gender differences with respect to retired tennis matches.

Acknowledgment

The authors would like to thank the section-editor and anonymous reviewers for their valuable comments and suggestions. This work has been financed by the Slovenian Research Agency within the EURO-CORES Programme EUROGIGA (project GReGAS) of the European Science Foundation and the project L5-2365.

References

- Barnett, T., Brown, A. and Pollard, G. (2006) Reducing the likelihood of long tennis matches. *Journal of Sports Science and Medicine* **5**, 567-574.
- Bastholt, P. (2000) Professional tennis (ATP Tour) and number of medical treatments in relation to type of surface. *Medicine and Science in Tennis* **5**(2). Available from URL: <http://www.stms.nl>
- Batagelj, V. and Mrvar, A. (1996-2011) *Pajek-Program for Large Network Analysis*. Available from URL: <http://pajek.imfm.si/doku.php?id=download>
- Batagelj, V. and Zaveršnik, M. (2011) Fast algorithms for determining (generalized) core groups in social networks. *Advances in Data Analysis and Classification* **5**, 129-145.
- Carrington, P. J., Scott, J. and Wasserman, S. (2005) *Models and Methods in Social Network Analysis*. Cambridge: Cambridge University Press.
- Clausen, S. E. (1998) *Applied Correspondence Analysis: An Introduction (Quantitative Applications in the Social Sciences)*. London: Sage Publications, Inc.
- Del Coral, J. (2009) Competitive balance and match uncertainty in grand-slam tennis: effects of seeding system, gender, and court surface. *Journal of Sports Economics* **10**, 563-581.
- Del Coral, J. and Prieto-Rodriguez, J. (2010) Are differences in ranks good predictors for Grand Slam tennis matches? *International Journal of Forecasting* **26**, 551-563.
- De Nooy, W., Mrvar, A. and Batagelj, V. (2011) *Exploratory Social Network Analysis with Pajek*. Revised and Expanded Second Edition. Cambridge: Cambridge University Press.
- ESPN. (2007) *Tennis officials investigate irregular betting on match*. Available from <http://sports.espn.go.com/sports/tennis/news/story?id=2959748>
- Fernandez, J., Mendez-Villanueva, A. and Pluim, B. M. (2006) Intensity of tennis match play. *British Journal of Sports Medicine* **40**, 387-391.
- Fernandez-Fernandez, J., Kinner, V. and Ferrauti, A. (2010) The physiological demands of hitting and running in tennis on different surfaces. *Journal of Strength and Conditioning Research* **24**, 3255-3264.
- Geyer, H. (2010) Quit behavior of professional tennis players. *Sports Medicine* **11**, 89-99.
- Girard, O., Eicher, F., Fouchet, F., Micallef, J.P. and Millet, G.P. (2007) Effects of the playing surface on plantar pressures and potential injuries in tennis. *British Journal of Sports Medicine* **41**, 733-738.
- Girard, O., Micallef, J. P. and Millet, G.P. (2010) Effects of the playing surface on plantar pressures during the first serve in tennis. *International Journal of Sports Physiology and Performance* **5**, 384-393.
- Girard, O. and Millet, G. P. (2004) Effects of the ground surface on the physiological and technical responses in young tennis players. In: *Science and racket sports*. Ed: Lees, A., Kahn, J.-F. and Maynard, I. W. London: Routledge.
- Greenacre, M. (2007) *Correspondence Analysis in Practice*. 2nd edition. Boca Raton: Chapman and Hall/CRC.
- Koning, R. H. (2010) Home advantage in professional tennis. *Journal of Sports Sciences* **29**, 19-27.
- Kovacs, M. S. (2006) Hydration and temperature in tennis - a practical review. *Journal of Sports Science and Medicine* **5**, 1-9.
- Magal, M., Webster, M.J., Sistrunk, L.E., Whitehead, M.T., Evans, J.C. and Boyd, R.K. (2003) Comparison of glycerol and water hydration regimens on tennis-related performance. *Medicine and Science in Sports and Exercise* **35**, 150-156.

- Murias, J.M., Lanatta, D., Arcuri, C.R. and Laino, F.A. (2007) Metabolic and functional responses playing tennis on different surfaces. *The Journal of Strength and Conditioning Research* **21**, 112-117.
- Nigg, B.M. and Segesser, B. (1988) The influence of playing surfaces on the load on the locomotor system and on football and tennis injuries. *Sports Medicine* **5**, 375-385.
- O'Donoghue, P. and Ingram, B. (2001) A notational analysis of elite tennis strategy. *Journal of Sports Sciences* **19**, 107-115.
- Pluim, B.M., Miller, S., Dines, D., Renstrom, P., Windler, G., Norris, B., Stroia, K.A., Donaldson, A. and Martin, K. (2007) Sport science and medicine in tennis. *British Journal of Sports Medicine* **41**, 703-704.
- Pluim, B.M., Staal, J.B., Windler, G.E. and Jayanthi, N. (2006) Tennis injuries: occurrence, aetiology, and prevention. *British Journal of Sports Medicine* **40**, 415-423.
- Pollard, G., Cross, R. and Meyer, D. (2006) An analysis of ten years of the four Grand Slam men's singles data for lack of independence of set outcomes. *Journal of Sports Science and Medicine* **5**, 561-566.
- R Development Core Team. (2011) *R: A language and environment for statistical computing* [Computer software manual]. Vienna, Austria. Available from URL: <http://www.R-project.org/>
- Wasserman, S. and Faust, K. (1994) *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

Key points

- The proportion of retired matches among professional tennis players has been increasing recently.
- Clay and hard courts are the most risky surfaces in relation to retired matches, particularly if the match is played at an outdoor venue.
- The difference in rankings of both players is proportional to the number/proportion of retired matches in professional tennis.
- Network analytic techniques could serve as an effective method to ascertain (a) group(s) of tennis players with the highest number of retired matches played among them.

AUTHORS BIOGRAPHY



Kristijan BREZNIK

Employment

Associate Professor, International School for Social and Business Studies, Celje, Slovenia

Degree

MSc

Research interests

Sport statistics, network analysis

E-mail: kristijan.breznik@mfdps.si



Vladimir BATAGELJ

Employment

Professor, Faculty of mathematics and Physics, University of Ljubljana, Slovenia

Degree

PhD

Research interests

Graph theory, network analysis

E-mail: vladimir.batagelj@fmf.uni-lj.si

✉ **Kristijan Breznik**

International School for Social and Business Studies, Mariborska 7, SI-3000 Celje, Slovenia