Int. j. racket sports sci. vol. 4(2), 2022, 40-55. eISSN: 2695-4508
DOI: 10.30827/Digibug.--- Original Investigation

Performance analysis in tennis since 2000: A systematic review focused on the methods of data collection

Análisis del rendimiento en tenis desde el año 2000: una revisión sistemática enfocada en los métodos de recolección de datos

Hiroo Takahashi [1] *, Shuhei Okamura [2] and Shunsuke Murakami [3]

1 Professor, National Institute of Fitness and Sports in Kanoya, Kagoshima, Japan.
2 Assistant Professor, Osaka University of Health and Sports Sciences, Osaka, Japan.
3 Senior Assistant Professor, National Institute of Fitness and Sports in Kanoya, Kagoshima, Japan.

Received: 01-04-22

Accepted: 03-03-23

Abstract

In tennis, performance analysis has advanced primarily as notational analysis. And analytical techniques markedly advanced, particularly in the fields of notational analysis and match analysis. In tennis, the Hawk-Eye system was introduced to tour tournaments in 2002. It has recently become used for player tracking and post-match analysis, there are a number of papers using Hawk-Eye data. Along with the development such measuring devices, technologies for analysis of a vast amount of data collected with these devices (big data) have also been developed. In particular, analysis by machine learning using AI was developed in the field of engineering, and it is also increasingly adopted in the field of sports. In the present review, we aimed to clarify the direction of research on performance analysis of tennis by organizing the trend of studies of performance analysis after 2000 with particular attention to the methods of data collection in the hope of furthering the development of this field. As a result of search of reports concerning performance analysis of tennis published after 2000 with particular interest in data collection methods, 90 papers were retrieved. The data collection methods were classified into active and passive methods, and subclassified into categories, i.e., tracking, video recording, data mining, observations of coaches, websites, and broadcasting. This review of the papers in different categories may aid in developing future directions of research in the field of performance analysis in tennis.

Keywords: *tracking, coding system, video recording, public data, simulation.*

Resumen

En tenis, el análisis del desempeño ha evolucionado principalmente como análisis notacional. Y las técnicas analíticas han avanzado de manera notable, especialmente en los campos del análisis notacional y de partidos. En tenis, el sistema Hawk-Eye fue incorporado a los torneos de circuito en 2002. Recientemente se ha usado para el seguimiento de jugadores y el análisis posterior al partido, y existen diversos artículos que usan datos del Hawk-Eye. Junto con el desarrollo de dichos dispositivos de medición, también se ha desarrollado tecnología para el análisis de grandes cantidades de datos recolectados con estos dispositivos (macrodatos). En particular, se desarrolló en el campo de la ingeniería el análisis con aprendizaje automático e IA, y cada vez es más usado en el ámbito deportivo. En esta revisión, el objetivo fue clarificar la dirección de la investigación sobre el análisis del rendimiento en tenis al organizar la tendencia de los estudios de análisis del rendimiento después del año 2000 con particular atención a los métodos de recolección de datos con el fin de continuar con el desarrollo de este campo. Como resultado de la búsqueda de artículos relacionados con el análisis del rendimiento en tenis publicados después del año 2000 enfocada en métodos de recolección de datos, se encontraron 90 artículos. Los métodos de recolección de datos fueron clasificados en activos y pasivos, y subclasificados en categorías, por ejemplo, seguimiento, grabación de video, minado de datos, observaciones de entrenadores, sitios web y transmisiones. Esta revisión de artículos en diferentes categorías puede ayudar en el desarrollo de otras líneas de investigación futuras en el campo del análisis del rendimiento en tenis.

Palabras clave: *seguimiento, sistema de codificación, grabación de video, datos públicos, simulación.*

Corresponding author: Hiroo Takahashi, hiroo@nifs-k.ac.jp

Cite this article as:
Jindo, T., Mitsuhashi, D., & Kubota, T. (2022). Accuracy of subjective stats of key performance indicators in tennis. *International Journal of Racket Sports Science,* 4(2), 40-55.

INTRODUCTION

Performance analysis is a new concept. Lees (2003) reviewed studies on racket sports by field and selected notational analysis as a category, but did not mention performance analysis.

The paper by Bartlett (2001) is considered to be the first on performance analysis. On the basis of differences between earlier biomechanical studies and studies of notational analysis, Bartlett (2001) defined the value of performance analysis as analysis of good and bad performances of the team and players according to performance indicators used in each genre of studies.

Thereafter, O'Donoghue (2010) defined performance analysis as investigation of sports performance using analytical methods that not only include biomechanics and notational analysis as reported by Bartlett, but also target data collected by physiological and psychological techniques.

In tennis, performance analysis has advanced primarily as notational analysis. As objectives of notational analysis, Hughes (1998) mentioned 1) tactical evaluation; 2) technical evaluation; 3) analysis of movement; 4) development of a database and modelling; and 5) educational use for both coaches and players, and reviews have since been reported according to these 5 goals. O'Donoghue (2004) also wrote reviews using the 5 goals proposed by Hughes (1998), but suggested, as prospects for the future, transformation of match analysis itself with the development of its techniques in addition to the necessity of conducting practical match analysis in the context of coaching.

The development of analytical techniques was previously brought up by Liebermann et al. (2002). Describing analytical methods for sports performance using the latest IT technology at the time, Liebermann et al. (2002) proposed that these technologies should be utilized in everyday coaching.

These reviews generally targeted papers published before 2000. Thereafter, analytical techniques markedly advanced, particularly in the fields of notational analysis and match analysis. In tennis, the Hawk-Eye system was introduced to tour tournaments in 2002 (hawkeyeinnovations.com, online). The initial objective of this system was to assist line judges, but as it has recently become used for player tracking and post-match analysis, there are a number of papers using Hawk-Eye data . In addition, instruments for tracking of the ball and players, such as Trackman (Trackman Inc.) and PlaySight (PlaySight Interactive ltd.), have been developed, and studies using such instruments are being conducted (Edelmann-Nusser et al., 2019; Murata and Takahashi, 2020; Kashiwagi et al., 2021).

Along with the development such measuring devices, technologies for analysis of a vast amount of data collected with these devices (big data) have also been developed. In particular, analysis by machine learning using AI was developed in the field of engineering, and it is also increasingly adopted in the field of sports. In studies concerning tennis, machine learning has been used by researchers including Whiteside and Reid (2017), Ganser et al. (2021), and Fernandes (2017).

In view of the changes in the methods for collection and analysis of data related to performance analysis of tennis, we considered it necessary to evaluate research themes of future performance analysis of tennis based on a review of papers published after 2000, when such changes became apparent.

In the present review, we aimed to clarify the direction of research on performance analysis of tennis by organizing the trend of studies of performance analysis after 2000 with particular attention to the methods of data collection in the hope of furthering the development of this field.

METHODS

This review was conducted according to the procedure of systematic review (Pickering and Byrne, 2014). To retrieve the relevant literature, searches were performed with "tennis", "performance", "analysis", "notation", and "match" as search words with the 'AND' condition, excluding "table" and "paddle" to restrict the search to reports concerning tennis. The databases searched were PubMed, Web of Science, and SPORT DISCUS, which encompass the literature concerning sports science. Searches were performed using the above search words in the default mode of each database, by which papers were retrieved if the search words were included in the title, abstract, or keywords. Two additional conditions, i.e., in English and published after 2000, were used for the search. The last date that we searched was April 23rd, 2021.

By the above method, 1,068 papers were retrieved. Of the retrieved papers, those that fulfilled the following conditions were included in the present review: 1) studies of performance in matches, 2) studies aiming to develop analytical methods, 3) studies analyzing quantitative data, and 4) studies published in the category of "research paper" in each journal. Papers that corresponded to the following were excluded as irrelevant to the objective of the present review: 1) studies focusing on physiological, psychological, and/or biomechanical indices alone as analytical targets, and 2) studies focusing on techniques of tennis and/or their development alone. In the first step, all the papers were screened by title, and all the authors agreed on the 130 papers that were retrieved. These papers were screened by Abstract to identify those that fulfilled the inclusion criteria, and all the authors agreed on the 90 papers that were retrieved.

While classifying the retrieved literature, attention was paid to the methods of data collection employed in each study. After overviewing the 90 retrieved papers, they were classified according to the data

collection method from the following viewpoints: 1) primary data collection: match data collected using videos and tracking systems at the sites of actual matches or data collected by the researchers themselves using audio-visual media, and analytical data prepared by the researchers themselves by conducting simulations using data from such sources, and 2) secondary data collection: data collected from broad sources, such as those made public online, or that were broadcast on the television. In addition, reports classified into 1) and 2) were subclassified according to the data collection method, and the characteristics of the subclasses were evaluated. The procedures of present review were showed on Figure 1.

RESULTS

According to the data collection methods, 42 and 48 of the 90 papers were classified as using primary and secondary data collection methods, respectively.

1) Studies using primary data collection

Primary data collection methods were subclassified into automatic data collection using tracking technologies, collecting data from video images, handling data from data mining, and collecting data from the observations of coaches. Tracking technologies were classified into vision-based technologies and inertial measurement unit (IMU)-based technologies. Data from video images were classified into automatic and manual methods.

1-1) Studies using tracking technologies

There were 16 studies using vision-based tracking technologies (Table 1-1 and 1-2). In this category, the study using Hawk-Eye data by Loffing et al. (2010) was published earliest. Of the reports evaluated in this review, all those using tracking technologies were published after 2010. Publication of studies using Hawk-Eye data increased, particularly after about 2016 (Kolbinger and Lames, 2013; Mecheri et al., 2016; Reid et al., 2016; Wei et al., 2016; Kovalchik and Albert, 2017; Kovalchik and Reid, 2018; Cui et al., 2019; Meurs et al., 2021). A characteristic of these studies was a large data size. Among the studies using Hawk-Eye data, the study by Mecheri et al. (2016) collected data from more than 100,000 points, and the study by Kovalchik and Albert (2017) targeted more than 30,000 services. They can be regarded as big data analyses.

There were two studies using IMU-based technologies (Table 1-2). A study by Myers et al. (2019) adopted Sony Smart Sensor, and a study by Edelmann-Nusser et al. (2019) adopted BABOLAT and HEAD sensors. However, as these studies reported that the measurement accuracy of the sensors were inacceptable, the studies that used these sensors were not published.

The reports in the table are arranged in: 1) chronological order, and 2) alphabetical order using the name of the authors. Studies by the same authors and those using the same methods are unified in the same row.

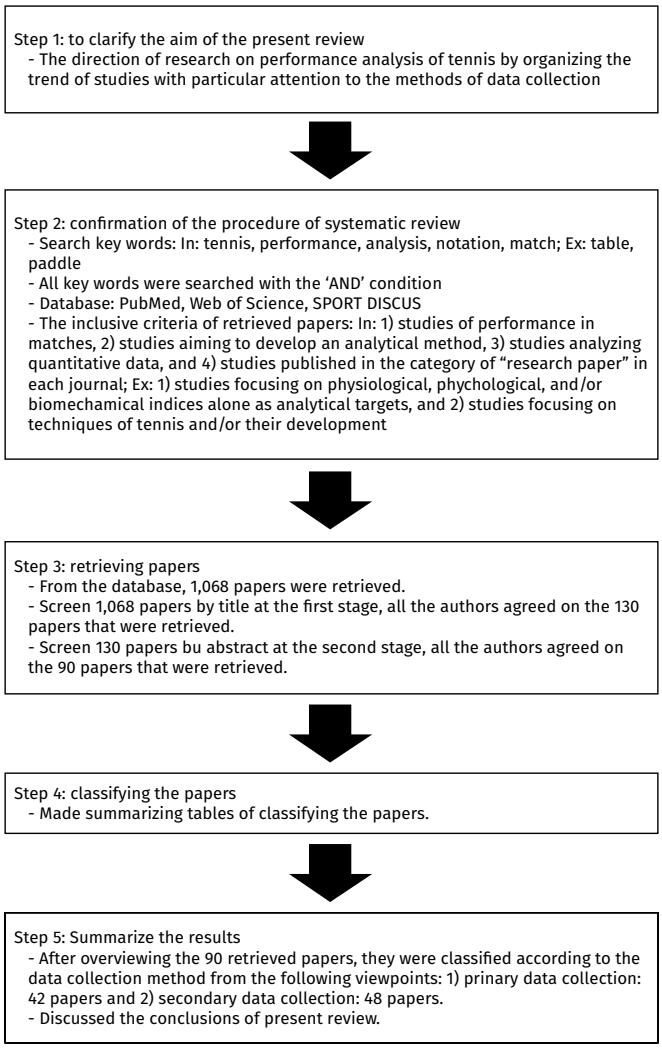

Figure 1. Score sheet.

1-2) Studies using video images

There were three studies that collected video images by the automatic method (Table 2-1). These studies used an independently developed system that processed video images automatically.

There were 18 studies that collected video images by the manual method (Table 2-2). These studies had been published since 2000. The methods of data collection in this category consisted of two types: observation of video images (Johnson and McHugh, 2006; Jans, 2007; Mergheş et al., 2014; Schmidhofer et al., 2014; Martin-Lorente et al., 2017), and developing independent systems (Klaassen and Magnus, 2003; Hizan et al., 2010; 2011; 2014; 2015; Klaus et al., 2017; Prieto-Lage et al., 2018). Many of studies targeted singles matches from Grand Slam tournaments.

Table 1-1. Data collection methods of vision-based tracking studies.

Authors	Year	Subject	Methods	Output data
Loffing et al.	2010	8098 rallies from 37 men's and 17 women's matches played at ATP, WTA, and Grand Slam	Hawk-Eye	% of the ball placements on opponent's backhand side
Kolbinger and Lames	2013	10418 serves of 53 right-handed male players from 56 men's singles Grand Slam matches of 2010 and 2011on hard court	Hawk-Eye	the placement of the ball of right-handed men's serves
Martínez-Gallego et al. Martínez-Gallego et al. Martinez-Gallego et al.	2013a 2013b 2019	188 games in 8 matches recorded at the ATP tournament 500 Valencia in 2011 11 professional players (age 24.8 ± 2.9) ranked between 5 and 113 on the ATP ranking	the SAGIT tracking system	(2013a) distance covered, average speed, time spent in the areas (2013b) % of unforced errors, % of winners and forced errors (2019) time, distance coveres, speed, winners, errors
Stare et al.	2015	boys U14 (n=11) and girls U14 (n=10) in the national championships in Slovenia ATP tournaments (n=7)	the SAGIT tracking system	the efficiency of the first and second serves the efficiency of the forehand and backhand the efficiency of the forehand and backhand in the return of serve the efficiency of topspin forehand or backhand, the slice of forehand or backhand
Mecheri et al.	2016	professional tennis tournaments (ATP and WTA) including Grand Slam between 2003 and 2008 75587 points for the women 187009 points for the men	Hawk-Eye	the relationships between the various characteristics of the serve (speed, location, spin, etc) and winning-point probabilities
Reid et al.	2016	102 male and 95 female players during the 2012-2014 Australian Open	Hawk-Eye	Serve performance Return of serve performance Groundstroke performance Movement characteristics
Wei et al.	2016	8780 shots of the top 3 players (Djokovic, Nadal, Federer) in the 2012 men's Australian Open	Hawk-Eye	Ground stroke speed ratio Ground stroke depth ratio Ground stroke angle ratio Lateral player movement ratio
Kovalchik and Albert	2017	175 matches from 2016 Australian open 87 matches of men and 88 matches of women	Hawk-Eye	time-to-serve rally length shot importance
Pereira et al.	2017	8 professional players during 4 matches of an international tournament (Futures level) on outdoor clay court in Brazil	Automatic tracking system by Figueroa et al. (2006)	Physical performance Technical performance
Kovalchik and Reid	2018	246 matches and 270,023 shots from men and 257 matches and 178,136 shots from women in 2015-2017 Australian open	Hawk-Eye	shot types (clustered by location, shape and speed) % of point won
Pereira et al.	2018	10 of U18 players from ITF tournament 8 professional players from Futures 10 professional players from ATP250	Automatic tracking system by Figueroa et al. (2006)	Time spent of interpersonal coordination patterns during lateral displacements: Anti-phase, In-phase, Serving player phase and Returning player phase
Cui et al.	2019	1188 of men, 189 individual players, from four Grand Slam's 2015-2017	Hawk-Eye	technical-tactical and physical performance
Floyd et al.	2020	5 matches from 2015 US Open	no show (only showed as 'tennis player-tracking data')	ESV (Expected Shot Value)
Meurs et al.	2021	64 men's matches from 2017 Australian open	Hawk-Eye	PA (Positional Advantage) index by Carvalho et al. (2013)

Table 1-2. Data collection methods of IMU-based tracking studies.

Authors	Year	Subject	Methods	Output data
Myers et al.	2019	14 junior players, 12 males and 2 females	Sony Smart Tennis Sensor	Hitting volume Ball speed
Edelmann-Nusser et al.	2019	4 matches by 8 players (10-18yrs, 4 female, 4 male) 2,098 strokes	BABOLAT PURE DRIVE PLAY BABOLAT POP HEAD Tennis Sensor PlaySight	number of strokes service speed

Table 2-1. Data collection methods of video images by automatic system studies.

Authors	Year	Subject	Methods	Output data
Connaghan et al.	2013	twelve complete matches with players of various skill levels, 825 min in total	Automated tennis event indexing system	accuracy of event detection
		same as above	Match Point: visual coding system	user's evaluation
Polk et al.	2014	two-set match of the best singles players on the coaches' team	TennisVis	the scoreline by Pie Meter View point outcome by Fish Grid View match summary by Filters and Bar Charts
Lara et al.	2018	a simulated match by two players	comparison of the manual versus automatic tracking	player's positioning

1-3) Studies using data mining

There were two studies using data mining theory (Table 3). These studies aimed to predict the results of matches or simulate the progression of matches.

1-4) Studies using data from the observations of coaches

A study that aimed to clarify performance analysis in tennis using data of the observations of coaches (Torres-Luque et al., 2018) was classified into this category (Table 4).

2) Studies of secondary data collection

Secondary data collection methods were subclassified into data collection from official websites and data collection from video images published by television broadcasting and websites.

2-1) Studies using data collected from websites

There were 38 studies using information released on websites (Table 5). Such studies were more common after 2010. Most of the studies targeted men's singles matches and collected data from the official ATP website and official Grand Slam website. Some studies targeted women's, doubles, and junior matches (Brenzik, 2013; Kovalchik et al., 2017; Cui et al., 2018; Sogut, 2018; Fernandez-Garcia et al., 2021; Li et al., 2020; Grambow et al., 2021). Other studies collected data from websites that gathered match data independently (Kovalchik and Reid, 2017; Kovalchik and Ingram, 2018; Fagan et al., 2019; Ingram, 2019; Makino et al., 2020). In addition, there were some studies that had no information about the data source (Pollard et al., 2006; Newton and Aslam, 2009; Tudor et al., 2014; Gu and Saaty, 2019; Stefani, 2020). A characteristic of these studies was their large data size.

2-2) Studies using data collected from broadcasting

There were 10 studies using data collected from broadcasting (Table 6). The studies collecting data from terrestrial and satellite broadcasting were published between 2000 and 2012 (O'Donoghue, 2001; O'Donoghue and Ingram, 2001; Gillet et al., 2009; Yu et al., 2009; Nowak and Panfil, 2012), whereas recent studies collected video images from websites (Carboch et al., 2018a, 2018b, 2019, 2020; Martinez-Gallego et al., 2020). Most of the studies targeted singles matches of Grand Slam tournaments, and one study targeted doubles matches (Martinez-Gallego et al., 2020).

DISCUSSION

1) Studies using primary data collection

Methods using automated vision-based tracking techniques, mainly Hawk-Eye, will continue to be the mainstay of primary data collection. Concerning studies using Hawk-Eye data, a group participated in by Tennis Australia has recently been active in reporting Australian Open matches (Reid et al., 2016; Wei et al., 2016; Kovalchik and Albert, 2017; Kovalchik and Reid, 2018; Meurs et al., 2021). There have also been studies focusing on other tournaments (Loffing et al., 2010; Kolbinger and Lames, 2013; Mecheri et al., 2016; Cui et al., 2019) and expansion of the research field was confirmed. The Hawk-Eye system is routinely employed in major tournaments. Groups conducting these studies reached an agreement with the tournament organizers about the use of the data obtained in the tournaments by the Hawk-Eye system for research. Building such relationships between tournament organizers and Hawk-Eye providers is considered a process indispensable for the development of research in this field.

Table 2-2. Data collection methods of video images by manual system studies.

Authors	Year	Subject	Methods	Output data
Klaassen and Magnus	2001 2003	(2001) 481 matches (male: 258, female: 223) at Wimbledon during 1992-1995	(2001,2009) no information for data collection	(2001) dynamic binary panel data with random effects
Klaassen and Magnus	2009	57,319 points in male, 28,979 points in female (2003,2009) all singles matches at Wimbledon 1992-1995	(In each match we know the two players and the complete sequence of points.) (2003) TENNISPROB	tests whether points in professional tennis are iid (independent and identically distributed) (2003) forecasting the probability of winning a match (2009) the efficiency of winning a point on serve
Johnson and McHugh	2006	22 players on 3 Grand Slams (8 in RG, 11 in Wimbledon, 9 in US) in 2003	observation from video recording	number of strokes stroke distribution
Jans	2007	3 final matches from 3 Grand Slams (RG, Wimbledon, US) in 2005	observation from video recording	time duration of point time interval of point total time of match time of play
Hizan et al.	2010	tennis coding system	coded the same match on two occasions separated by a 4-week period 5 raters coded 674 shots	intra-rater reliability inter-rater reliability comparison with Hawk-Eye data
Hizan et al. Hizan et al. Hizan et al.	2011 2014 2015	(2011) 28 matches (male:14, female: 14) from 2008 Australian Open, 2666 points (male: 1651, female: 1015) 28 U-16 (male: 14, female: 14) matches and 28 U-12 (male: 14, female: 4) matches from 2008 Australian Boys and Girls championships, 2359 points on U-16 (male: 1239, female: 1120) and 2267 points on U-12 (male: 1175, female: 1092) (2014) 23 matches (male:11, female: 12) from 2008 Australian Open, 1968 successful serves (male:1172, female: 796) 27 U-16 (male: 14, female: 13) matches and 21 U-12 (male: 12, female: 9) matches from 2008 Australian Boys and Girls championships, 2836 succesful serves on U-16 (male: 1439, female: 1397) and 1647 succesful serves on U-12 (male: 916, female: 731) (2015) 23 matches (male:11, female: 12) from 2008 Australian Open, 5221 serves (male: 3272, female: 1949) 27 U-16 (male: 14, female: 13) matches and 21 U-12 (male: 12, female: 9) matches from 2008 Australian Boys and Girls championships, 3391 serves on U-16 (male: 1740, female: 1651) and 1922 serves on U-12 (male: 1050, female: 872)	tennis coding system (by Hizan et al., 2010)	(2011) % 1st in, aces, DF, % 1st won, % 2nd won, % 1st return won, % 2nd return won (2014) serve-return location point winning (2015) serve location point winning
Carvalho et al. Carvalho et al.	2013 2014	(2013) 27 rallies in 3 matches from 2008 Estoril Open (ATP 250) (2014) 28 rallies in 3 matches from 2008 Estoril Open (ATP 250)	recording by DV camera and 2D-DLT	(2013) PA (Positional Advantage) index (2014) GDD (Goal-Directed Displacement) index
Mergheş et al.	2014	9 matches by 3 players (Federer, Nadal, Agassi) in 2 years	observation from video recording	% won on 1st serve % won on 2nd serve % won on return

Table 2-2. Data collection methods of video images by manual system studies (Continuation).

Schmidhofer et al.	2014	12 matches each for 3 groups (U9, U10, U12) from Develop Tournaments for Australian Tennis Association 12 matches of ATP tournaments	observation from video recording	service parameters return parameters ICT (Inter Contact Times) parameters miscellaneous parameters
Fitzpatrick et al.	2017	48 participants MTR: n=18, Age 7.4 ± 0.6, 230 points MTO: n=16, Age 8.5 ± 0.6, 253 points MTG: n=8, Age 9.9 ± 0.4, 280 points FB: n=6, Age 13.7 ± 0.5 247 points	a custom-notational analysis system	service parameters return parameters ICT (Inter Contact Times) parameters miscellaneous parameters
Klaus et al.	2017	8 U-14 national level male players in Australia QF and SF of the Victorian Junior Hardcourt Championships	A developed computerized system Kinovea (version 0.8.15)	type of stroke type of outcome court position
Martin-Lorente et al.	2017	18 matches of Grand Slam and ATP finals between 2011 and 2014 11 men players	observation from video recording	results of inside out and inside in forehand
Prieto-Lage et al.	2018	82 break point events between Nadal and Djokovic on final clay court during 2011 and 2012	observation from video recording with OBSTENNIS	the break points T-Pattern
Martínez-Gallego et al.	2021	2339 points from 19 complete doubles matches of the 2018 ATP World Tour Masters 1000 tournament played in Canada	a data collection system was designed using Microsoft Excel	time characteristics of doubles tennis time characteristics of the points by winning and losing team time characteristics of the points by the type of match

Table 3. Data collection methods of data mining studies.

Authors	Year	Subject	Methods	Output data
O'Donoghue and Simmonds	2019	Traditional tennis games Traditional tiebreaks Fast4 tennis games Tiebreaks in Fast4 tennis Tiebreak Ten	Simulation in various winning point probabilities	The probability of the player who serve first
Li et al.	2021	no information	data mining technology	serve points won and lost

Table 4. Data collection methods of studies using the observations of coaches.

Authors	Year	Subject	Methods	Output data
Torres-Luque et al.	2018	observational instrument	video observation by one observer questionnaire to 10 experts	the list of variables and categories related with the result of the match the list of variables and categories related with the development of the game

Table 5. Data collection methods of public data on websites.

Authors	Year	Subject	Source	Output data
Pollard et al.	2006	4883 matches data from 1995-2004 All Grand Slam tournaments of men's singles	no show	the probability of winning a set in a match iid (independent and identically distributed) in a set
Djurovic et al.	2009	128 matches data from 2007 and 2008 Grand Slam hard court tournaments from IBM	IBM DB2 applicaction	the latent (factor) area of a tennis match
Newton and Aslam	2009	330 players over 59 ATP tournaments in 2007 four Grand Slams four top players (Federer, Nadal, Roddick, Blake)	no show	percentage of points won on serve percentage of points won on receiving serve Monte Carlo simulations
Reid et al.	2010	2007 Matchfact information of the top 100 male professional players	ATP website	correlation coefficients between the different performance variables
O'Donoghue	2012	Study 1: 92 men's singles matches in the 2011 US Open from the official website Study 2: world top four players in Grand Slam tournament between 2008 and 2011	Study 1: 2011 US Open website Study 2: official Grand Slam website	expected and observed break points per receiving game probability of winning points during receiving game
Breznik	2013	male (N=16,732) and female (N=16,432) players between 1968 to 2011 obtained from ATP and WTA website	ATP website WTA website	number of matches won by handedness results of PageRank algorithm
Ma et al.	2013	18,288 performances between 1991 and 2008 from the website of the ATP	ATP website	predicting winner or loser by logistic regression model with three variables (match characteristics, personal characteristics, skills and performance)
O'Donoghue	2013	men's and women's matches from 2012 Grand Slam tournaments	official Grand Slam website	propotion of points won on serving probability of rare events occurred
Vaverka and Cernosek	2013	players participated in all four Grand Slam in 2008	official yearbooks and website of the ITF official Grand Slam website	correlation coefficients between body height and serve speed
Bane et al.	2014	rankings data and date of birth information from 1985 to 2010	ATP website	the age of first ATP ranking the time to reach Top 100 from first ranked the time between first entry and exit from Top 100 the time between first ATP ranked and exit from Top 100
Kovalchik	2014	498 competitors in end-of-year ATP rankings of 104 or higher between 1991 and 2012	ATP website	trends in player characteristics (30 and over, teenagers, the age of peak performance, etc) with local polynomial regression curves
Tudor et al.	2014	all the matches from main draws of Roland Garross, Wimbledon and US Open in 2010 and 2011	no show	match statistics
Filipcic et al.	2015	male players ranked in top 300 on ATP ranking in 1991, 2000 and 2010 match statistics of 1961 matches from 1991, 2363 matches from 2000, 2660 matches from 2010	ATP website	match statistics
Kim et al.	2015	2012 Australian Open SF video on the web a men's match and a women's match	video on a website (site information is no show) A coordinate system by Matlab	location of the ball bounce time series of ball anble differences
Kovalchik	2016	53,442 matches played by ATP top 100 players in 2004-2014 and 1,377 matches from 2015	ATP website	fitted model of the Pythagorean theorem

Table 5. Data collection methods of public data on websites (Continuation).

Prieto-Bermejo et al.	2016	ATP top 10 players on four Grand Slams between 1990 and 2012	ATP website	relationships between ranking position and the results on tournaments
Kovalchik and Reid	2017	match activity from 2000-2015 of junior players from ITF website professional men's and women's players from Tennis Abstract website point-by-point data for Grand Slam matches from FlashScore website more detailed point-by-point data from Hawk-Eye	ITF website Tennis Abstract website FlashScore website Hawk-Eye	relative importance of match statistics for winning
Kovalchik et al.	2017	877 player trajectories entered WTA rankings between 1989 and 2016	WTA website	the mean peak ranking in the first ranking year, the number of years during which the majority of progression occurred (the progression stage), and the rate of rankings gained during the progression stage
Cui et al.	2017	1188 players in 594 matches collected from four 2015-2017 Grand Slams men's singles	official Grand Slam website	relationships between match statistics and relative quality (RQ)
Cui et al.	2020b			difference of performance indicators between seeded players and non-seeded players
Cui et al.	2018	1369 matches in four Grand Slams women's singles	official Grand Slam website	between match variables and the relative quality (RQ) performance profiles
Kovalchik and Ingram	2018	1582 men's matches and 966 women's matches from 2010 to the present 33,788 points across 161 men's matches and 21,450 points across 170 women's matches at the 2015 and 2016 Australian Opens by Hawk-Eye	the Match Charting Project (www.tennisabstract.com) Hawk-Eye	point distribution by match format time distribution by match format impact of match format on match durations and upsets
Sogut	2018	male (n=60) and female (n=59) players in 2017 Wimbledon	ATP website WTA website	correlation between body height and mtatch outcomes
Vaverka et al.	2018	men (n=72-92) and women (n=70-98) at four Grand Slams in 2008, 2012 and 2016	official Grand Slam website	differences in the serve speed of Grand Slams
Fagan et al.	2019	handedness data as well as match- play results from ATP Tennis in 2014	ATP Tennis Navigator (http://www.tennisnavigator.com/)	the advantage of left-handedness probability of match-play results
Fitzpatrick et al.	2019a	244 men's matches and 250 women's matches from 2016 and 2017 French Open	2016 and 2017 Roland Garros website	relationships between performance characteristics and PWOL (Percentage of matches in which the Winner Outscored the Loser)
Fitzpatrick et al.	2019b	244 men's matches and 250 women's matches from 2016 and 2017 French Open 241 men's matches and 249 women's matches from 2016 and 2017and Wimbledon	Roland Garros website and the Wimbledon information System by IBM	relationships between performance characteristics and PWOL (Percentage of matches in which the Winner Outscored the Loser)
Gu and Saaty	2019	82987 matches from 1990 for ATP and 35886 matches from 2003 for WTA	online sites	predicted the outcome of 2015 US OPEN
Ingram	2019	2208 matches from ATP 2014 season	MatchStat.com (scraping)	a point-based Bayesian hierarchical model for predicting the outcome of tennis matches (the probability of winning a point on serve given surface, tournament and match date)
Martin et al.	2019	50 five-set matches from 2014 Grand Slams	official Grand Slam website	effect of pacing strategies on match outcome effect of players' ATP ranking on pacing strategies effect of Grand Slam tournament on pacing strategies
Cui et al.	2020a	146 men's matches from 2016-2017 US Open and Australian Open	official website of each tournament	set-to-set differences of match performance

Table 5. Data collection methods of public data on websites (Continuation).

Damani et al.	2020	127 men's matches from 2020 Australian Open	2020 Australian Open website	differences of match statistics among entire tournament, initial rounds (1R-4R) and intense rounds (QF, SF and F)
Fernandez-Garcia et al.	2020	546 matches by professionals and U-18 in three Grand Slams	official Grand Slam (Australian Open, Roland Garros and Wimbledon) website	differences of match statistics between professionals and U-18 players
Grambow et al.	2020	1772 men's matches from 2002-2015 Wimbledon	Wimbledon information System by IBM	serve performance comparisons by tournaments year and tournament week
Li et al.	2020	professional players of mens (n=180) and womens (n=193) within top 300 ranking between 2010 and 2018	ATP website WTA website	relationships between the age and their ranking milestones
Makino et al.	2020	4230 points on three surfaces (Hard, Clay, Grass) of four players (Federer, Nadal, Murray, Djokovic)	Match Charting Project (https://github.com/JeffSackmann/tennis_MatchChartingProject)	match winner predictions using machine learning
Stefani et al.	2020	almost 5000 men's and 5000 women's matches of four Grand Slams from 2006-2019	no show	percent of matches by the higher-seeded players
Grambow et al.	2021	1771 ladies' matches from 2002-2015 Wimbledon	Wimbledon information System by IBM	serve performance comparisons by tournaments year and tournament week

Table 6. Data collection methods of broadcasting studies.

Authors	Year	Subject	Methods	Output data
O'Donoghue	2001	men's and women's 252 matches from Grand Slam tounaments between 1997 and 1999 from terrestrial and satellite television coverage	a computerized data management system	proportion of points won when serving proportion of games won
O'Donoghue and Ingram	2001	men's and women's 175 matches from Grand Slam tounaments between 1997 and 1999 from terrestrial and satellite television coverage	a specially designed computerized notational analysis system for tennis	differencese of timing factors and strategy data among tournaments and gender
Gillet et al.	2009	116 men's matches from French Grand Slam tournament in 2005 and 2006 from terrestrial television coverage	a computerized notational system	serve characteristics and point winning serve-return characteristics and point winning
Yu et al.	2009	broadcast tennis video	a frame grouping technique	3D virtual content insertion application ball detection and tracking application
Nowak and Panfil	2012	the match by Federer and Djokovic of 2007 US Open final and 2008 Australian Open semi-final from broadcasts by Eurosport	data recorded with Microsoft Excel	relationships among type of shot, ball placement on court and fixed or dynamic elements of play
Carboch et al. 2018a Carboch et al. 2018b Carboch et al. 2019		23 women's matches from 2017 Australian Open 7 men's and 23 women's matches from 2017 Australian Open 24 men's matches from Austrarian Open, French Open and Wimbledon in 2017 from television or internet broadcast	a spreadsheet for observed variables	comparisons of point duration, number of rally shots, time between the points, rally pace and work to rest ratio
Carboch et al.	2020	23 women's matches from 2017 Australian Open and 24 men's matches from Austrarian Open, French Open and Wimbledon in 2017 from television or internet broadcast	a spreadsheet for observed variables	comparisons of match characteristics between new and used balls
Martínez-Gallego et al.	2020	34 men's doubles matches from ATP tournaments in 2018 from Tennistv.com	a registration system created with Microsoft Excel	point ending situations

As mentioned above, there was a limitation to the use of the Hawk-Eye system data; thus, video images obtained by the manual method were used. The manual method of collecting video images was a general methodology. Especially the studies that targeted junior matches, such as those without the Hawk-Eye system, adopted the manual method to collect video images (Schmidhofer et al., 2014; Fitzpatrick et al., 2017; Klaus et al., 2017). Recently, advances in image processing have made it easier to calculate parameters from images obtained with video cameras than before 2000. As the use of video images is a relatively simple method to collect data in environments where it is difficult to employ a high-tech system, such as Hawk-Eye, the use of methods currently employed in other sports events and image processing techniques used in other fields as well as developing original systems for automatic collection of parameters from video images appropriate for the objective of the study using existing techniques as references may be solutions for the establishment of a method for data collection from video images.

We confirmed that two studies used data mining theory (O'Donoghue and Simmonds, 2019; Li et al., 2021). As mentioned below, there were many studies that used published data on the Internet. The field of data mining was prospected to develop a technique for predicting or simulating the results of matches with published data on the Internet.

2) Studies of secondary data collection

Many studies of secondary data collection were carried out by collecting data from websites. On the present website of the ATP Tour (ATP TOUR.com, online), a wide variety of data, including the summary of points scored and the decisive shot at each score called MATCH BEATS, detailed results of rallies called RALLY ANALYSIS, and, on the page called the second-screen, positions where the ball was hit, positions where the ball fell, distance run, and speed of the ball hit, in addition to conventional stats, such as the first-service percentage and first-service scoring rate, are provided. Such detailed data has the same quality as the vision-based tracking data described in this review, and proceeding with exploratory research using such open data may lead to further development of research in the field of performance analysis in tennis. In particular, many studies analyzing such data from a long-time perspective have been conducted, and they are expected to provide findings that will aid in the 4) development of a database and modelling, and 5) educational use for both coaches and players among the 5 viewpoints suggested by Hughes (1998) by making studies from both cross-sectional and longitudinal viewpoints possible.

However, public data from tournaments and matches are limited, and only data of particular tournaments are available. In addition, it was only after 1991 that stats began to be provided and after 2018 that detailed stats began to be released. Therefore, caution is needed in the use of data.

Recently, data collected from broadcasting have become available on websites as streaming services. Data collected by such methods will continue to be used for research.

Most of the studies by secondary data collection targeted men's singles matches of world top-ranked players. There were few studies of female players, doubles matches, and junior players. It is necessary to perform studies to obtain data about these categories. As mentioned below, studies that targeted junior matches, such as those that were played without the Hawk-Eye system, adopted the manual method to collect video images, especially data from online streaming video for doubles matches.

CONCLUSIONS

As a result of search of reports concerning performance analysis of tennis published after 2000 with particular interest in data collection methods, 90 papers were retrieved. The data collection methods were classified into primary and secondary methods, and subclassified into 6 categories, i.e., tracking, video recording, data mining, the observations of coaches, Internet, and broadcasting. This review of the studies in different categories suggests the importance of considering vision-based tracking technologies, the increased use of manual video-recordings, the possibility of data mining, the use of official websites, and performing studies focusing on female players, doubles teams, and junior players.

REFERENCES

ATP TOUR.com (online) Nitto ATP Finals | Results | ATP Tour | Tennis. https://www.atptour.com/en/scores/current/nitto-atp-finals/605/results?, (accessed 2021-11-17).

Bane, M., Reid, M., Morgan, S. (2014). Has player development in men's tennis really changed? An historical rankings perspective. *Journal of Sports Sciences, 32*(15), 1477-1484.

Bartlett, R. (2001). Performance analysis: can bringing together biomechanics and notational analysis benefit coaches? *International Journal of Performance Analysis in Sport, 1*(1), 122-126. https://doi.org/10.1080/24748668.2001.11868254

Breznik, K. (2013). On the gender effects of handedness in professional tennis. *Journal of Sports Science and Medicine, 12*(2), 346-353.

Carboch, J., Blau, M., Sklenarik, M., Siman, J., Placha, K. (2020). Ball change in tennis: How does it affect match characteristics and rally pace in Grand Slam tournaments? *Journal of Human Sport and Exercise, 15*(1), 153-162. https://doi.org/10.14198/jhse.2020.151.14

Carboch, J., Placha, K. (2018a). Development of rally pace and other match characteristics in women's matches in the Australian open 2017. *Journal of Physical Education and Sport, 18*(2), 1079-1083. https://doi.org/10.7752/jpes.2018.s2161

Carboch, J., Placha, K., Sklenarik, M. (2018b). Rally pace and match characteristics of male and female tennis matches at the Australian Open 2017. *Journal of Human Sport and Exercise, 13*(4), 743-751. https://doi.org/10.14198/jhse.2018.134.03

Carboch, J., Siman, J., Sklenarik, M., Blau, M. (2019). Match characteristics and rally pace of male tennis matches in three grand slam tournaments. *Physical Activity Review, 7*, 49-56. https://doi.org/10.16926/par.2019.07.06

Carvalho, J., Araújo, D., Travassos, B., Esteves, Pedro., Pessanha, L., Pereira, F., Davids, K. (2013). Dynamics of players' relative positioning during baseline rallies in tennis. *Journal of Sports Sciences, 31*(14), 1596-1605. https://doi.org/10.1080/02640414.2013.792944

Carvalho, J., Araújo, D., Travassos, B., Fernandes, O., Pereira, F., Davids, K. (2014). Interpersonal Dynamics in Baseline Rallies in Tennis. *International Journal of Sports Science & Coaching, 9*(5), 1043-1056. https://doi.org/10.1260/1747-9541.9.5.1043

Connaghan, D., Moran, K., O'Connor, NE. (2013). An automatic visual analysis system for tennis. *Proceedings of The Institution of Mechanical Engineers Part P-Journal of Sports Engineering and Technology, 227*(4), 273-288. https://doi.org/10.1177/1754337112469330

Cui, Y., Gomez, MA., Goncalves, B., Liu, H., Sampaio, J. (2017). Effects of experience and relative quality in tennis match performance during four Grand Slams. *International Journal of Performance Analysis in Sport, 17*(5), 783-801. https://doi.org/10.1080/24748668.2017.1399325

Cui Y, Gómez MÁ, Gonçalves B, Sampaio J. (2018). Performance profiles of professional female tennis players in grand slams. *PLOS One, 13*(7), e0200591. https://doi.org/10.1371/journal.pone.0200591

Cui, Y., Gomez, MA., Goncalves, B., Sampaio, J. (2019). Clustering tennis players' anthropometric and individual features helps to reveal performance fingerprints. *European Journal of Sport Science, 19*(8), 1032-1044. https://doi.org/10.1080/17461391.2019.1577494

Cui, Y., Liu, H., Gómez, MÁ., Liu, H., Gonçalves, B. (2020a). Set-to-set Performance Variation in Tennis Grand Slams: Play with Consistency and Risks. *Journal of Human Kinetics, 73*(1), 153-163. https://doi.org/10.2478/hukin-2019-0140

Cui Y, Zhao Y, Liu H, Gómez MÁ, Wei R, Liu Y. (2020b). Effect of a seeding system on competitive performance of elite players during major tennis tournaments. *Frontiers in Psychology, 11*, 1294. https://doi.org/10.3389/fpsyg.2020.01294

Damani, C., Damani, B., Bagchi, A. (2020). Match statistics significant to win the initial and intense rounds of a tennis tournament. *TRENDS in Sport Sciences, 27*(4), 225-231. https://doi.org/10.23829/TSS.2020.27.4-6

Djurovic, N., Lozovina, V., Pavicic, L. (2009). Evaluation of tennis match data - new acquisition model. *Journal of Human Kinetics, 21*, 15-21. https://doi.org/10.2478/v10078-09-0002-9

Edelmann-Nusser, A., Raschke, A., Bentz, A., Montenbruck, S., Edelmann-Nusser, J., Lames, M. (2019). Validation of sensor-based game analysis tools in tennis. *International Journal of Computer Science in Sport, 18*(2), 49-59. https://doi.org/10.2478/ijcss-2019-0013

Fagan, F., Haugh, M., Cooper, Hal. (2019). The advantage of lefties in one-on-one sports. *Journal of Quantitative Analysis in Sports, 15*(1), 1-25. https://doi.org/10.1515/jqas-2017-0076

Fernandes, M. A. (2017). Using Soft Computing Techniques for Prediction of Winners in Tennis Matches. *Machine Learning Research, 2*(3), 86-98. https://doi.org/10.11648/j.mlr.20170203.12

Fernández-García, A. I., Giménez-Egido, J. M., and Torres-Luque, G. (2021). Differences in Grand Slam competition statistics between professional and U-18 players according to the sex. *Revista Internacional de Ciencias del Deporte, 17*(63), 25-37. https://doi.org/10.5232/ricyde2021.06303

Figueroa, P. J., Leite, N. J., & Barros, R. M. L. (2006). Tracking soccer players aiming their kinematical motion analysis. *Computer Vision and Image Understanding, 101*(2), 122–135. https://doi.org/10.1016/j.cviu.2005.07.006

Filipcic, A., Zecic, M., Reid, M., Crespo, M., Panjan, A., Nejc, S. (2015). Differences in performance indicators of elite tennis players in the period 1991-2010. *Journal of Physical Education & Sport, 15*(4), 671-677. https://doi.org/10.7752/jpes.2015.04102

Fitzpatrick, A., Davids, K., Stone, JA. (2017). Effects of Lawn Tennis Association mini tennis as task constraints on children's match-play characteristics. *Journal of Sports Sciences, 35*(22), 2204-2210. https://doi.org/10.1080/02640414.2016.1261179

Fitzpatrick, A., Stone, JA., Choppin, S., Kelley, J. (2019a). A simple new method for identifying performance characteristics associated with success in elite tennis. *International Journal of Sports Science & Coaching, 14*(1), 43-50. https://doi.org/10.1177/1747954118809089

Fitzpatrick, A., Stone, JA., Choppin, S., Kelley, J. (2019b). Important performance characteristics in elite clay and grass court tennis match-play. *International Journal of Performance Analysis in Sport, 19*(6), 942-952. https://doi.org/10.1080/24748668.2019.1685804

Floyd, CM., Hoffman, M., Fokoue, E. (2020). Shot-by-shot stochastic modeling of individual tennis points. *Journal of Quantitative Analysis in Sports, 16*(1), 57-71. https://doi.org/10.1515/jqas-2018-0036

Ganser, A., Hollaus, B. and Stabinger, S. (2021). Classification of Tennis Shots with a Neural Network Approach. *Sensors*, *21*(17), 5703. https://doi.org/10.3390/s21175703

Gillet E, Leroy D, Thouvarecq R, Stein JF. (2009). A notational analysis of elite tennis serve and serve-return strategies on slow surface. *Journal of Strength Conditioning and Research, 23*(2), 532-539. https://doi.org/10.1519/JSC.0b013e31818efe29

Grambow, R., O'shannessy, C., Born, P., Meffert, D., Vogt, T. (2020). Serve efficiency development at Wimbledon between 2002 and 2015: A longitudinal approach to impact tomorrow's tennis practice. *Human Movement, 21*(1), 65-72. https://doi.org/10.5114/hm.2020.88155

Grambow, R., O'Shannessy, C., Born, P., Meffert, D., Vogt, T. (2021). Serve efficiency development indicates an extended women's tennis world class cohort: Analysing 14 years of Ladies Wimbledon Championships - implications for coaching. *Human Movement, 22*(2), 43-52. https://doi.org/10.5114/hm.2021.100011

Gu, W., Saaty, TL. (2019). Predicting the outcome of a tennis tournament: based on both data and judgments. *Journal of Systems Science and Systems Engineering, 28*(3), 317-343. https://doi.org/10.1007/s11518-018-5395-3

hawkeyeinnovations.com (online) Hawk-Eye innovations. https://resources.platform.pulselive.com/HawkEye/document/2016/07/29/ec84be34-2375-4b5a-8854-917e0e7021f0/HawkEyeinTennis2016.pdf, (accessed 2021-10-8).

Hizan, H., Whipp, P., Reid, M. (2010). Validation of Match Notation (A Coding System) in Tennis. *Journal of Quantitative Analysis in Sports, 6*(3), 1-13, https://doi.org/10.2202/1559-0410.1223

Hizan, H., Whipp, P., Reid, M. (2011). Comparison of serve and serve return statistics of high performance male and female tennis players from different age-groups. *International Journal of Performance Analysis in Sport, 11*(2), 365-375. https://doi.org/10.1080/24748668.2011.11868556

Hizan, H., Whipp, P., Reid, M. (2015). Gender differences in the spatial distributions of the tennis serve. *International Journal of Sports Science & Coaching, 10*(1), 87-96. https://doi.org/10.1260/1747-9541.10.1.87

Hizan, H., Whipp, P., Reid, M; Wheat, J. (2014). A comparative analysis of the spatial distributions of the serve return. *International Journal of Performance Analysis in Sport, 14*(3), 884-893, https://doi.org/10.1080/24748668.2014.11868765

Hughes, M. (1998). The application of notational analysis to racket sports. In Science and Racket Sports II (edited by Lees, A., Maynard, I., Hughes, M. and Reilly, T.), pp211-220. London: E & FN Spon.

Hughes, M. and Franks, I. (2004a). Notational analysis – a review of the literature. In Notational Analysis of Sport Second edition (edited by Hughes, M. and Franks, I.), pp59-106. London: Routledge.

Hughes, M. and Franks, I. (2004b). From analysis to coaching. In Notational Analysis of Sport Second edition (edited by Hughes, M. and Franks, I.), pp257-271. London: Routledge.

Ingram, M. (2019). A point-based Bayesian hierarchical model to predict the outcome of tennis matches. *Journal of Quantitative Analysis in Sports, 15*(4), 313-325. https://doi.org/10.1515/jqas-2018-0008

Jans, W. (2007). The Time Structure of Female Tennis Players During the Grand Slam Tournaments. *Research Yearbook, 13*(2), 230-234.

Johnson, CD., McHugh, MP. (2006). Performance demands of professional male tennis players. *British Journal of Sports Medicine, 40*(8), 696-699. https://doi.org/10.1136/bjsm.2005.021253

Kashiwagi, R., Okamura, S., Iwanaga, S., Murakami, S., Numata, K. and Takahashi, H. (2021). The differences in the ball speed and the spin rate depending on the results of a tennis serve. *Malaysian Journal of Movement, Health & Exercises, 10*(1), 48-50. https://doi.org/10.4103/2231-9409.328217

Kim, H., Cai, F., Ryu, J., Haddad, JM., Zelaznik, HN. (2015). Tennis match time Series do not exhibit long term correlations. *International Journal of Sport Psychology, 46*(6), 542-554. https://doi.org/10.7352/IJSP.2015.46.542

Klaassen, FJGM., Magnus, JR. (2001). Are points in tennis independent and identically distributed? Evidence from a dynamic binary panel data model. *Journal of the American Statistical Association, 96*(454), 500-509. https://doi.org/10.1198/016214501753168217

Klaassen, FJGM., Magnus, JR. (2003). Forecasting the winner of a tennis match. *European Journal of Operational Research, 148*(2), 257-267. https://doi.org/10.1016/S0377-2217(02)00682-3

Klaassen, FJGM., Magnus, JR. (2009). The efficiency of top agents: An analysis through service strategy in tennis. *Journal of Econometrics, 148*(1), 72-85. https://doi.org/10.1016/j.jeconom.2008.09.036

Klaus, A., Bradshaw, R., Young, W., O'Brien, B., Zois, J. (2017). Success in national level junior tennis: Tactical perspectives. *International Journal of Sports Science & Coaching, 12*(5), 618-622. https://doi.org/10.1177/1747954117727792

Kolbinger, O., Lames, M. (2013). Ball trajectories in tennis - Lateral and vertical placement of right handed men's singles serves. *International Journal of Performance Analysis in Sport, 13*(3), 750-758. https://doi.org/10.1080/24748668.2013.11868686

Kovalchik, SA. (2014). The older they rise the younger they fall: age and performance trends in men's professional tennis from 1991 to 2012. *Journal of Quantitative Analysis in Sports, 10*(2), 99-107. https://doi.org/10.1515/jqas-2013-0091

Kovalchik, SA. (2016). Is there a Pythagorean theorem for winning in tennis? *Journal of Quantitative Analysis in Sports, 12*(1), 43-49. https://doi.org/10.1515/jqas-2015-0057

Kovalchik, SA., Albert, J. (2017). A multilevel Bayesian approach for modeling the time-to-serve in professional tennis. *Journal of Quantitative Analysis in Sports, 13*(2), 49-62. https://doi.org/10.1515/jqas-2016-0091

Kovalchik, SA., Reid, M. (2017). Comparing matchplay characteristics and physical demands of junior and professional tennis athletes in the era of big data. *Journal of Sports Science and Medicine, 16*(4), 489-497.

Kovalchik, S., Bane, M., Reid, M. (2017). Getting to the top: an analysis of 25 years of career rankings trajectories for professional women's tennis. *Journal of Sports Sciences, 35*(19), 1-7. https://doi.org/10.1080/02640414.2016.1241419

Kovalchik, SA., Ingram, M. (2018). Estimating the duration of professional tennis matches for varying formats. *Journal of Quantitative Analysis in Sports, 14*(1), 13-23. https://doi.org/10.1515/jqas-2017-0077

Kovalchik, S., Reid, M. (2018). A shot taxonomy in the era of tracking data in professional tennis. *Journal of Sports Sciences, 36*(18), 2096-2104. https://doi.org/10.1080/02640414.2018.1438094

Lara, JPR., Vieira, CLR., Misuta, MS., Moura, FA., de Barros, RML. (2018). Validation of a video-based system for automatic tracking of tennis players. *International Journal of Performance Analysis in Sport, 18*(1), 137-150. https://doi.org/10.1080/24748668.2018.1456886

Lees, A. (2003). Science and the major racket sports: a review. *Journal of Sports Sciences, 21*, 707-732.

Li, M., Li, Q., Li, Y., Cui, Y., Zhao, X., Guo, L. (2021). Analysis of characteristics of tennis singles matches based on 5G and data mining technology. *Security and Communication Networks*, 5549309. https://doi.org/10.1155/2021/5549309

Li, P., Weissensteiner, JR., Pion, J., Bosscher, VD. (2020). Predicting elite success: Evidence comparing the career pathways of top 10 to 300 professional tennis players. *International Journal of Sports Science & Coaching, 15*, 793-802. https://doi.org/10.1177/1747954120935828

Liebermann, D. G., Katz, L., Hughes, M. D., Bartlett, R. M., McClements, J. and Franks, I. M. (2002). Advances in the application of information technology to sports performance. *Journal of Sports Sciences, 20*(10), 755-769. https://doi.org/10.1080/026404102320675611

Loffing, F., Hagemann, N., Strauss, B. (2010). Automated processes in tennis: do left-handed players benefit from the tactical preferences of their opponents? *Journal of Sports Sciences, 28*(4), 435-443. https://doi.org/10.1080/02640410903536459

Ma, SM., Liu, CC., Tan, Y., Ma, SC. (2013). Winning matches in Grand Slam men's singles: An analysis of player performance-related variables from 1991 to 2008. *Journal of Sports Sciences, 31*(11), 1147-1155. https://doi.org/10.1080/02640414.2013.775472

Makino, M., Odaka, T., Kuroiwa, J., Suwa, I., Shirai, H. (2020). Feature selection to win the point of ATP tennis players using rally information. *International Journal of Computer Science in Sport, 19*(1), 37-50. https://doi.org/10.2478/ijcss-2020-0003

Martin-Lorente, E., Campos, J., Crespo, M. (2017). The inside out forehand as a tactical pattern in men's professional tennis. *International Journal of Performance Analysis in Sport, 17*(4), 429-441. https://doi.org/10.1080/24748668.2017.1349528

Martin, C., Bideau, B., Touzard, P., Kulpa, R. (2019). Identification of serve pacing strategies during five-set tennis matches. *International Journal of Sports Science & Coaching, 14*(1), 32-42. https://doi.org/10.1080/24748668.2017.1349528

Martínez-Gallego, R., Crespo, M., Ramón-Llin, J., Micó, S., Guzmán, JF. (2020). Men's doubles professional tennis on hard courts: Game structure and point ending characteristics. *Journal of Human Sport & Exercise, 15*(3), 633-642. https://doi.org/10.14198/jhse.2020.153.13

Martinez-Gallego, R., Guzman, JF., Crespo, M., Ramon-Llin, J., Vuckovic, G. (2019). Technical, tactical and movement analysis of men's professional tennis on hard courts. *Journal of Sports Medicine and Physical Fitness, 59*(1), 50-56. https://doi.org/10.23736/S0022-4707.17.07916-6

Martinez-Gallego, R., Guzman, JF., James, N., Pers, J., Ramon-Llin, J., Vuckovic, G. (2013a). Movement characteristics of elite tennis players on hard courts with respect to the direction of ground strokes. *Journal of Sports Science and Medicine, 12*(2), 275-281.

Martínez-Gallego, R., Guzmán, JF., James, N., Ramón-Llin, J., Crespo, M., Vuckovic, G. (2013b). The relationship between the incidence of winners/errors and the time spent in different areas of the court in elite tennis. *Journal of Human Sport & Exercise, 8*(3), S601-607, https://doi.org/10.4100/jhse.2013.8.Proc3.05

Martínez-Gallego, R., Vives, F., Guzmán, JF., Ramón-Llin, J., Crespo, M. (2021). Time structure in men's professional doubles tennis: does team experience allow finishing the points faster? *International Journal of Performance Analysis in Sport, 21*(2), 215-225. https://doi.org/10.1080/24748668.2021.1872218

Mecheri, S., Rioult, F., Mantel, B., Kauffmann, F., Benguigui, N. (2016). The serve impact in tennis: First large-scale study of big Hawk-Eye data. *Statistical Analysis and Data Mining: The ASA Data Science Journal, 9*, 310-325. https://doi.org/10.1002/sam.11316

Mergheş, PE., Simion, B., Nagel, A. (2014). Comparative analysis of return of serve as counter-attack in modern

tennis. *Timisoara Physical Education & Rehabilitation Journal, 6*(12), 18-22. https://doi.org/10.2478/tperj-2014-0023

Meurs, EV., Buszard, T., Kovalchik, S., Farrow, D., Reid, M. (2021). Interpersonal coordination in tennis: assessing the positional advantage index with Australian Open Hawkeye data. *International Journal of Performance Analysis in Sport, 21*(1), 22-32. https://doi.org/10.1080/24748668.2020.1843213

Murata, M. and Takahashi, H. (2020). Verification of the accuracy and reliability of the TrackMan tennis radar. Proceedings of the Institution of Mechanical Engineers, Part P: *Journal of Sports Engineering and Technology. 235*(2), 154-160. https://doi.org/10.1177/1754337120953005

Myers, NL., Kibler, WB., Axtell, A.H., Uhl, TL. (2019). The Sony Smart Tennis Sensor accurately measures external workload in junior tennis players. *International Journal of Sports Science & Coaching, 14*(1), 24-31. https://doi.org/10.1177/1747954118805278

Newton, P., Aslam, K. (2009). Monte Carlo Tennis: A Stochastic Markov Chain Model. *Journal of Quantitative Analysis in Sports, 5*(3), 7. https://doi.org/10.2202/1559-0410.1169

Nowak, M., Panfil, R. (2012). Scoring abilities in the game of tennis - a pragmatic study of unique cases. *Human Movement, 13*(4), 313-322. https://doi.org/10.2478/v10038-012-0036-z

O'Donoghue, PG. (2001). The most important points in Grand Slam singles tennis. *Research Quarterly for Exercise and Sport, 72*(2), 125-131. https://doi.org/10.1080/02701367.2001.10608942

O'Donoghue, P. (2004). Match analysis in racket sports. Science and Racket Sports III (edited by Lees, A., Kahn, J. -F. and Maynard, I. W.), pp155-162. London: Routledge.

O'Donoghue, P. (2010). Research methods for sports performance analysis. London: Routledge, pp1-28.

O'Donoghue, P. (2012). Break points in Grand Slam men's singles tennis. *International Journal of Performance Analysis in Sport, 12*(1), 156-165. https://doi.org/10.1080/24748668.2012.11868591

O'Donoghue, P. (2013). Rare events in tennis. *International Journal of Performance Analysis in Sport, 13*(2), 535-552. https://doi.org/10.1080/24748668.2013.11868668

O'Donoghue, P., Girard, O. and Reid, M. (2013). Racket Sports. Routledge Handbook of Sports Performance Analysis (edited by McGarry, T., O'Donoghue, P. and Sampaio, J.), pp404-414. London: Routledge.

O'Donoghue, P., Ingram, B. (2001). A notational analysis of elite tennis strategy. *Journal of Sports Sciences, 19*(2), 107-115. https://doi.org/10.1080/026404101300036299

O'Donoghue, P., Simmonds, Emma. (2019). Probability of winning and match length in Tiebreak Ten tennis. *International Journal of Performance Analysis in Sport, 19*(3), 402-416. https://doi.org/10.1080/24748668.2019.1615296

Pereira, TJC., Nakamura, FY., de Jesus, MT., Vieira, CLR., Misuta, MS., de Barros, RML., Moura, FA. (2017). Analysis of the distances covered and technical actions performed by professional tennis players during official matches. *Journal of Sports Sciences, 35*(4), 361-368. https://doi.org/10.1080/02640414.2016.1165858

Pereira, TJC., van Emmerik, REA., Misuta, MS., Barros, RML, Moura, FA. (2018). Interpersonal coordination analysis of tennis players from different levels during official matches. *Journal of Biomechanics, 67*, 106-113. https://doi.org/10.1016/j.jbiomech.2017.11.036

Pickering, C. and Byrne, J. (2014). The benefits of publishing systematic quantitative literature reviews for PhD candidate and other early-career researchers. Higher Education Res Dev., 33(3): 534-548.

Polk, T., Yang, J., Hu, YQ., Zhao, Y. (2014). TenniVis: Visualization for tennis match analysis. *IEEE Transactions on Visualization and Computer Graphics, 20*(12), 2339-2348. https://doi.org/10.1109/TVCG.2014.2346445

Pollard, G., Cross, R., 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(4), 561-566.

Prieto-Bermejo, J., Gómez-Ruano, MÁ. (2016). Entering tennis men's Grand Slams within the top-10 and its relationship with the fact of winning the tournament. *Revista Internacional de Ciencias del Deporte, 12*(46), 411-422. https://dx.doi.org/10.5232/ricyde2016.04605

Prieto-Lage, I., Prieto, MA., Curran, TP., Gutierrez-Santiago, A. (2018). An accurate and rapid system to identify play patterns in tennis using video recording material: Break point situations as a case study. *Journal of Human Kinetics, 62*(1), 199-212. https://doi.org/10.1515/hukin-2017-0170

Reid, M., McMurtrie, D., Crespo, M. (2010). The relationship between match statistics and top 100 ranking in professional men's tennis. *International Journal of Performance Analysis in Sport, 10*(2), 131-138. https://doi.org/10.1080/24748668.2010.11868509

Reid, M., Morgan, S., Whiteside, D. (2016). Matchplay characteristics of Grand Slam tennis: implications for training and conditioning. *Journal of Sports Sciences, 34*(19), 1791-1798. https://doi.org/10.1080/02640414.2016.1139161

Schmidhofer, S., Leser, R., Ebert, M. (2014). A comparison between the structure in elite tennis and kids tennis on scaled courts (Tennis 10s). *International Journal of Performance Analysis in Sport, 14*(3), 829-840. https://doi.org/10.1080/24748668.2014.11868761

Sogut, M. (2018). Stature: Does it really make a difference in match-play outcomes among professional tennis

players? *International Journal of Performance Analysis in Sport,18*(2), 255-261. https://doi.org/10.1080/24748668.2018.1466259

Stare, M., Žibrat, U., Filipčič, A. (2015). Stroke effectiveness in professional and junior tennis. *Kinesiologia Slovenica, 21*(2), 39-50.

Stefani, R. (2020). A longitudinal analysis of the differential performances of seeded male and female Grand Slam tennis players. *The Sport Journal, 21*. https://thesportjournal.org/article/a-longitudinal-analysis-of-the-differential-performances-of-seeded-male-and-female-grand-slam-tennis-players/

Torres-Luque, G., Fernández-García, ÁI., Cabello-Manrique, D., Giménez-Egido, JM., Ortega-Toro, E. (2018). Design and validation of an observational instrument for the technical-tactical actions in singles tennis. *Frontiers in Psychology, 9*: 2418. https://doi.org/10.3389/fpsyg.2018.02418

Tudor, PB., Zecic, M., Matkovic, B. (2014). Differences between 2010 and 2011 performance indicators of tennis play at the grand slam tournaments. *Kinesiology, 46*, 101-106.

Vaverka, F., Cernosek, M. (2013). Association between body height and serve speed in elite tennis players. *Sports Biomechanics, 12*(1), 30-37. https://doi.org/10.1080/14763141.2012.670664

Vaverka, F., Nykodym, J., Hendl, J., Zhanel, J., Zahradnik, D. (2018). Association between serve speed and court surface in tennis. *International Journal of Performance Analysis in Sport, 18*(2), 262-272. https://doi.org/10.1080/24748668.2018.1467995

Wei, XY., Lucey, P., Morgan, S., Sridharan, S. (2016). Forecasting the next shot location in tennis using fine-grained spatiotemporal tracking data. *IEEE Transactions on Knowledge and Data Engineering, 28*(11), 2988-2997. https://doi.org/10.1109/TKDE.2016.2594787

Whiteside, D. and Reid, M. (2017). Spatial characteristics of professional tennis serves with implications for serving aces: A machine learning approach, *Journal of Sports Sciences, 35*:7, 648-654, https://doi.org/10.1080/02640414.2016.1183805

Yamaguchi, S., Oshimi, D. and Fukuhara, T. (2018). The impact of sport events on a host region: A literature review. Japan J. Phys. Educ. Hlth. *Sport Sci., 63*, 13-32. https://doi.org/10.5432/jjpehss.17065

Yu, XG., Jiang, NJ., Cheong, LF., Leong, HW., Yan, X. (2009). Automatic camera calibration of broadcast tennis video with applications to 3D virtual content insertion and ball detection and tracking. *Computer Vision and Image Understanding, 113*(5), 643-652. https://doi.org/10.1016/j.cviu.2008.01.006