Computational Statistics
https://doi.org/10.1007/s00180-024-01578-y

ORIGINAL PAPER

Eras of dominance: identifying strong and weak periods in professional tennis

Kristijan Breznik[1,2] · Vincenzo Candila[3] · Antonina Milekhina[4] · Marialuisa Restaino[3]

Received: 21 March 2024 / Accepted: 8 November 2024
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract
In sports journalism and among fans, there is an ongoing debate on identifying eras where the level of competition is extremely high. In tennis, a common question concerning the advent of the so-called 'Big Three'—listed alphabetically, Novak Djokovic, Roger Federer, and Rafael Nadal—is: Did these players lead to an unprecedented high level of competition? We contribute to this debate by identifying, from a statistical point of view, *strong* players, periods, and eras in men's tennis, where a *strong* era is defined as a time frame in which a subset of (*strong*) players consistently dominate all the others. Hence, this work extends the idea of the Greatest Player of All Time (GOAT), largely investigated in the literature, to a dynamic subset of players. Through cointegration analysis of over 30 years of professional tennis data, we identify five *strong* eras. Interestingly, the player with the largest participation during these *strong* eras is Roger Federer and the most recent strong era concluded in July 2019. Moreover, we examine the relationship between the match duration and *strong* players/periods/eras, finding that the occurrence of a match between *strong* and not-*strong* players decreases the match duration, on average. Furthermore, when *strong* players meet, the match duration generally increases.

Keywords Time series · OLS regression · Cointegration · ATP data · Match duration

1 Introduction

Over the last 30 years, the utilization of statistical tools in sports analysis has significantly increased (Schumaker et al. 2010; Morgulev et al. 2018; Baumer et al. 2023; Albert et al. 2017), thereby facilitating the development of predictive models for sports events, the assessment of player/team attributes, and the refinement of coaching tactics, besides other applications. Recently, the statistical methodologies have

Kristijan Breznik, Vincenzo Candila, Antonina Milekhina and Marialuisa Restaino have contributed equally to this work.

Extended author information available on the last page of the article

Published online: 20 November 2024

been used in several sports, like soccer (Kirschstein and Liebscher 2019; Alfano et al. 2021; Cefis and Carpita 2022; D'Urso et al. 2023; Bai et al. 2023), basketball (Sandri et al. 2020; Zuccolotto et al. 2023; de Paula Oliveira and Newell 2024), volleyball (George and Panagiotis 2008; Gabrio 2021; López-Serrano et al. 2024), and football (Baker and McHale 2013; Akhanli and Hennig 2023), among others.

Regarding individual sports, tennis plays a crucial role, being globally widespread, with competitions held on all continents almost continuously throughout the year and with approximately 87 million tennis players worldwide (Musa 2023a). The popularity of tennis has recently surged due to the presence of current and potentially future superstars like (listed alphabetically) Carlos Alcaraz, Novak Djokovic, Roger Federer, Rafael Nadal, and Jannik Sinner. For instance, the epic Wimbledon 2019 final between Federer and Djokovic attracted 9.6 million British viewers on BBC1 (Musa 2023b), while the final of the Association of Tennis Professionals (ATP) Finals 2023 between Djokovic and Sinner achieved an exceptional 29.5% share in the Italian television.

So far, different tennis-related topics have stimulated the application and development of specific statistical techniques in the literature, like network theory, machine learning tools, regression models, and pairwise comparisons. For instance, network analytic methods have been applied to measure the strength of tennis players in both singles competition (Radicchi 2011) and doubles (Breznik 2015). Dynamic network analysis was used to compare in-match psychological traits of tennis players by gender (Milekhina et al. 2023). Network centrality measures, specifically the eigenvector centrality measure, were employed to develop a model for predicting match winners (Arcagni et al. 2023). Machine learning techniques in tennis were also explored by Candila and Palazzo (2020) and Wilkens (2021), among others. Regression models appear to be one of the most popular techniques in tennis for forecasting purposes. In particular, regression models have been employed for predicting match winners (Del Corral and Prieto-Rodríguez 2010; Lisi and Zanella 2017; Gorgi et al. 2019), preventing sports injuries among tennis players (Wang and Yao 2024), analyzing susceptibility to pressure (Harris et al. 2021), and examining the effect of serve order in tennis tiebreaks (Cohen-Zada et al. 2018). Bayesian inference and models have been used to investigate intended serve direction for professional tennis players at Roland Garros (Tea and Swartz 2023), and to study player characteristics involving the ability to extend rallies, to pinpoint the differences between winning and losing, giving feedback on how players may improve (Dona et al. 2024). Bootstrapping is applied to estimate the points lost, games lost, sets lost and matches lost due to unforced errors (Peiris et al. 2024). Pairwise comparison approaches like the Elo Elo (1978), Kovalchik (2020) and Angelini et al. (2022) or the Bradley–Terry type model (McHale and Morton 2011) have been used to predict the outcome of the matches.

The topics and contributions mentioned earlier represent only a fraction of the scientific and statistically-based literature on tennis. However, the identification of a subset of players who consistently dominate all others for a significant period has not received adequate attention, leading to the ongoing debate, among sports journalists and fans, on *strong* and *weak* eras. In this respect, a question concerning the advent of the so-called 'Big Three'—listed alphabetically, Novak Djokovic, Roger

Federer, and Rafael Nadal—is: Did these players lead to an unprecedented high level of competition? This paper aims to contribute to this debate. Specifically, the objective of this study is to identify *strong* players, periods, and eras in men's tennis and examine their relationship with match duration. We define *strong* players as a subset who dominates all other players for a given period of time, a *strong* period as a time when at least two *strong* players are present, and a *strong* era as a time interval with at least two subsequent *strong* periods. We define *weak* period as a period where no *strong* players are present and *weak* era as a time interval with at least two subsequent *weak* periods. *Strong* and *weak periods* are mutually exclusive and collectively exhaustive phenomena. Sequences of *strong* and *weak periods* may or may not lead to *strong* and *weak eras*.

The identification of *strong* players, periods, and eras is carried out through the Engle–Granger two-step cointegration analysis (Engle and Granger 1987) of ATP points earned by the top-ranked players. We resort to the Engle-Granger approach because the order of the time series under investigation is important. While cointegration has been used in diverse fields beyond economic time series, such as pollution data (Ang 2007), tourism (Khan et al. 2005, and even football Dobson and Goddard (1996)), to the best of our knowledge, this study represents the first application of cointegration analysis in tennis. While the recognition of *strong* players/periods/eras is a novelty in tennis, other studies have investigated eras in baseball, hockey, and golf (Berry et al. 1999).

By identifying *strong* players and *strong/weak* periods/eras, we contribute to the literature in three significant aspects. Firstly, the identification of a subset of tennis players who, for a given amount of time, dominate all the other players can be seen as a contribution to the topic of the Greatest Player of All Time (GOAT), recently investigated by Radicchi (2011), Baker and McHale (2014) and Baker and McHale (2017) for the male and female circuits, respectively. The perspective expands beyond identifying a single GOAT by encompassing a cohort of players who exhibit sustained dominance over their rivals. This approach acknowledges that in sports, including emerging ones, initial dominance can set benchmarks for excellence even if subsequent players do not replicate the same level of dominance. Additionally, one's presence within a cohort of dominant players for some prolonged time provides another proof of one's greatness. Secondly, our approach for identifying *strong* players/periods/eras can be used to find statistically-based rivalries, which are key factors for enhancing fans' interest, improving tournament schedules, and capitalizing on specific marketing strategies, among other factors. Thirdly, we formulate and test two hypotheses concerning *strong* players and match duration: (i) *When a strong player faces a not-strong player, the match duration decreases, on average*; (ii) *When two strong players meet, the match duration increases, on average*. Examining the relationship between *strong* eras and match duration is of interest since evidence (Simmons 2006) suggests that prolonged durations negatively impact the demand for sports.

Analyzing the dataset of ATP points for the top ten positions from 1990 to 2023, we identified five *strong* eras. Interestingly, Roger Federer emerges as the player with the highest participation during these *strong* eras. The most recent *strong* era concluded in July 2019. Furthermore, considering a dataset of almost 87,000 matches, we present strong and robust evidence supporting the hypotheses on reduced match duration when a *strong* player meets a not-*strong* player, and increased match duration when *strong* players face each other. Therefore, there is evidence of the dominance of *strong* players who can win their matches quickly, on average. On the other hand, matches between two *strong* players tend to have increased match duration, on average and as expected, due to the closeness of the two contenders.

The rest of the paper is organized as follows. Section 2 illustrates the methodology used and introduces the details of both the definition of *strong* players/period/eras and the hypotheses. Section 3 is devoted to the empirical analysis, showing the main results. In particular, in Sect. 3.1, the *strong* players/periods/eras are shown for the analyzed time frame. Then, in Sect. 3.2, we test the two hypotheses to examine the relationship between match duration and the occurrence of *strong* players/periods/eras. In Sect. 3.3, a robustness analysis is performed to give consistency to the results of our analysis. Section 4 provides the discussion and conclusions.

2 Methodology and hypotheses

We aim at defining the notion of *strong* players and *strong* periods/eras and identifying such players, periods, and eras in male tennis.

Let i, t be the double time index for the match and week, respectively. The index i denotes the ith match of the week t, with $i = 1, \dots, N_t$ and $t = 1, \dots, T$, with T the total number of weeks. Let N_t be the number of matches contested in week t. Moreover, let $N = \sum_{t=1}^{T} N_t$ be the total number of matches.

To identify *strong* players and *strong* periods/eras, we examine the ATP points observed every week t for the players in the top positions of the ATP ranking. Therefore, the ATP points can be seen as a time series. Formally, let $y_{t,k}$ be the ATP points of the player in position k (for $k = 1, \dots, K$, where the K is the lowest position in the ranking) at week t. For instance, $y_{t,1}$ represents the ATP points that the player in position 1 has for the week t.

As a time series, the ATP points can be regarded as a sample realization of a stochastic process, which may or may not be stationary (Box et al. 2015). In the case of (weak) stationarity, the process exhibits a constant mean and variance over time, and an autocovariance that depends only on the lag considered, not on the time at which it is calculated. Conversely, in the case of non-stationarity, the process lacks a constant mean, a constant variance, or both, and the autocovariance may depend not only on the lag but also on the specific time at which it is calculated. One possible cause for the non-stationary is the presence of a stochastic trend or a structural break. Among the various statistical tests available in the literature to check stationarity, the Augmented Dickey and Fuller (ADF, 1979) test is one of the most used. The ADF test is designed to verify the presence of a unit root in the time series, implying the existence of a stochastic trend.

According to our notation, the ADF test is based on the following equation:

$$\Delta y_{t,k} = \alpha + \gamma y_{t-1,k} + \delta_1 \Delta y_{t-1,k} + \cdots + \delta_p \Delta y_{t-p,k} + u_t, \tag{1}$$

where Δ is the difference operator concerning the points of the player at position k, that is: $\Delta y_{t,k} = y_{t,k} - y_{t-1,k}$, p is the order of the lags adopted according to some information criteria, like the Akaike (AIC, Akaike 1974) or Bayesian (BIC, Schwarz 1978) criterion, $\alpha, \gamma, \delta_1, \ldots, \delta_p$ are the unknown coefficients, and u_t is the error term.

The hypotheses in the ADF test are:

$$\begin{cases} H_0 : \gamma = 0; \\ H_1 : \gamma < 0. \end{cases}$$

If $\gamma = 0$, it means that there is a unit root in the time series, which, consequently, is not stationary, while if $\gamma < 0$, the time series is stationary. Thus, if the null hypothesis fails to be rejected, the test provides evidence that the series is non-stationary.

However, when a linear combination of (two or more) non-stationary time series leads to a stationary series, the (two or more) series are said to be *cointegrated* (Engle and Granger 1987). In particular, two (or more) cointegrated time series share a stochastic trend, that is, there exists a stable long-run relationship between these time series.

Going back to the ATP points, our idea is that the ATP points of the player in position k, $y_{t,k}$, might be cointegrated with the ATP points of players at positions different from k.

The occurrence of cointegrated ATP points leads to the definition of the *strong players*:

Definition 1 (*Strong* players) A strong player is a top-ten player whose ATP points are cointegrated with at least another ATP top-ten player points, within a time window of length T_{in}.

In Definition 1, T_{in} represents the number of weeks used to identify the *strong* players. According to Definition 1, the search for *strong* players takes place via the following steps:

1. **Regress** the ATP points of the first player, $y_{t,1}$, on the ATP points of the players on lower positions, from the period $t = 1 + j$ to $t = T_{in} + j$.
2. **Employ** the ADF test on the residuals of the regression in the previous step.
3. **Iterate** Steps 1 and 2, with $j = \{0, 1, 2, \ldots\}$, until the end of the sample.

Step 1 uses the Ordinary Least Squares (OLS) method in the so-called "cointegrating" regression, that is:

$$y_{t,1} = \beta_0 + \beta_1 y_{t,2} + \beta_2 y_{t,3} + \cdots + u_t, \text{ with } t = 1 + j, \ldots, T_{in} + j. \tag{2}$$

Then, the cointegration test through Eq. (1) is performed on the regression residuals $\hat{u}_t$ obtained from Eq. (2). According to Hamilton (1994), the ADF test statistics resulting from the "cointegrating" regression in Step 1 are compared with the critical values derived from Phillips and Ouliaris (1990) and Hansen (1992). When the null hypothesis of non-stationarity is rejected, the ATP points of all the players included in the regression in Step 1 are cointegrated. Therefore, according to Definition 1, all the players occupying the positions involved in the regression during the period of length T_{in} are defined as *strong*.

Once the definition for the *strong* players has been established, we introduce the definitions of *strong* and *weak* periods/eras.

Definition 2 (*Strong* periods/eras) A strong period is a period when at least two strong players (according to Definition 1) are present.

A strong era is a time interval consisting of at least two subsequent strong periods.

Definition 3 (*Weak* periods/eras) A weak period is when no strong players (according to Definition 1) are present.

A weak era is a time interval consisting of at least two subsequent weak periods.

Therefore, every time the null hypothesis of the ADF test in Step 2 is rejected, the period under investigation is defined as *strong*. Otherwise, the period under investigation is referred to as a *weak* period. The eras are finally obtained, merging at least two subsequent (*weak* or *strong*) periods.

Identifying a *strong* era might lead to emerging high rivalries between strong players. Moreover, during *strong* periods/eras, the *strong* players generally dominate consistently over all other players. In contrast, during *weak* periods/eras, no group of players dominates all the others. Bearing this in mind, we formulate the following hypotheses concerning the relationship between the match duration and the occurrence of *strong* players:

Hypothesis 1 During strong periods, when a strong player faces a not-strong player, the match duration decreases, on average.

Hypothesis 2 When two strong players meet, the match duration increases, on average.

Hypothesis 1 suggests that during *strong* eras, characterized by the dominance of *strong* players, matches between *strong* and not-*strong* players may be shorter, on average, due to the disparity in skill levels. The high level of *strong* players can lead to faster wins.

Hypothesis 2 posits that when two strong players face each other, the match duration increases on average because the closely matched skill levels result in more competitive and prolonged exchanges. These encounters often feature players who are well-matched in terms of ability, leading to tighter games, more sets, and longer rallies, all of which contribute to longer matches.

Both hypotheses align with intuitive expectations about competitive dynamics and the nature of match play in professional tennis.

3 Empirical analysis

In this study, we merge two datasets, both freely available online at the GitHub project by Jeff Sackmann.

The first dataset provides information on the week-by-week occupier of each ranking place on the ATP tour and the related ATP points. The ATP points incorporate the results of the past 52 weeks (with some minor exceptions).

The second dataset concerns the day-by-day matches among professional players involved in Grand Slams, ATP Finals, ATP Tour Masters 1000, ATP Tour 500, and ATP Tour 250 tournaments. The dataset provides information about the winner, the loser, the final score, the match statistics (such as the overall duration expressed in minutes), and some characteristics of the tournament (for example, the type of surface, the level of the tournament, the draw, and the hosting city).

For both datasets, the sample period covers the time frame from January 1990 to December 2023. The total number of matches under analysis is $N = 86,823$.

Figure 1 shows the ATP points of the top 10 players for the period under investigation. Upon graphical inspection, a notable change in points was observed in 2009, where increased points are attributed to the final rounds of the most important ATP tournaments. Consequently, also the variability of points among the top ten players tends to be higher. Additionally, in 2020, points remained fixed due to the lockdowns implemented in many countries worldwide after the surge of Covid-19. It is worth noting that the points generally exhibit similar patterns, except for the points of the first ATP player (the top black line), particularly during the period from 2015 to 2017.

The results of ADF tests, according to the nine different periods considered, each of 4 years except the last, which is of 3 years, are reported in Table 1. For almost all periods, the null hypothesis of the unit root is not rejected, confirming our intuition that the ATP points are not stationary. Therefore, the next step is to

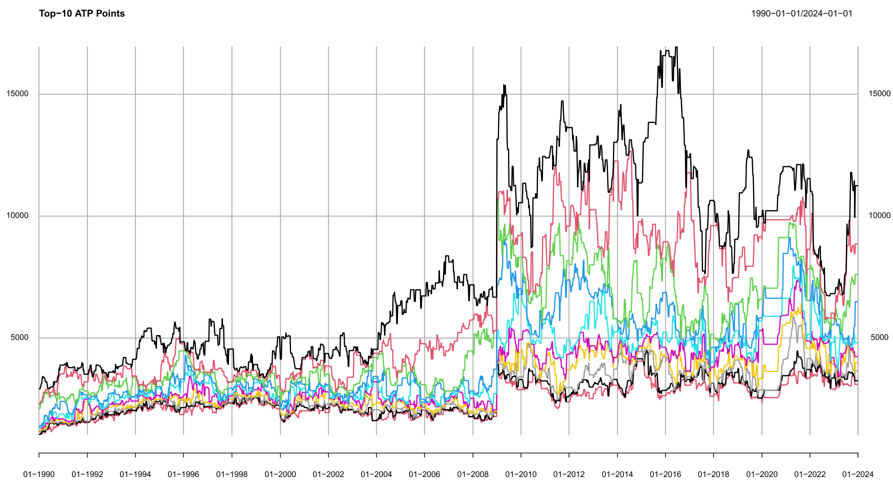

Fig. 1 The ATP points for the top-10 players from January 1990 to December 2023

Table 1 ADF test statistic for the ATP points of the players occupying the first ten positions (first column), according to nine different periods (first row)

Ranking	1990/1992	1993/1996	1997/2000	2001/2004	2005/2008	2009/2012	2013/2016	2017/2020	2021/2023
1	−2.271	−2.100	−2.178	−1.526	−2.216	−2.186	−1.391	−2.958	−0.411
2	−2.520	−1.731	−1.912	−2.250	−2.228	−2.139	−2.235	−2.665	−0.962
3	−2.576	−2.190	−2.242	−2.014	−1.822	−1.934	−2.092	−2.164	−0.209
4	−2.649	−2.931	−2.408	−−2.702	−2.406	−1.764	−1.803	−1.741	−0.394
5	−2.189	−2.967	−2.595	−2.459	−3.595***	−1.329	−2.924	−1.851	−2.649
6	−2.780	−3.779**	−1.871	−3.017	−3.272*	−2.531	−2.025	−1.981	−2.096
7	−2.955	−2.959	−1.478	−3.307*	−2.165	−2.060	−3.577**	−0.742	−2.012
8	−2.912	−2.795	−1.930	−4.187***	−2.132	−2.022	−2.634	−2.098	−3.245*
9	−3.107	−2.311	−2.054	−3.202*	−2.922	−2.366	−1.764	−1.850	−3.091
10	−3.208*	−2.791	−1.902	−3.168*	−2.292	−2.564	−1.794	−1.749	−3.333*

*, ** and *** represent the significance at levels 10%, 5%, 1%, respectively

Each period lasts 4 years, except the first and the last which last 3 years investigate the presence of *strong* players and periods/eras in Sect. 3.1 and test the Hypotheses 1 and 2 in Sect. 3.2.

3.1 Identification of *strong* players and periods/eras

In the first part of our analysis dedicated to the search for *strong* players (Definition 1), and subsequently *strong* and *weak* periods/eras (Definitions 2 and 3) through the three steps described above (in Sect. 2, on page 5), we initially set $T_{in} = 90$, meaning that the (smallest) potential *strong* period corresponds to 90 weeks. The number of regressors involved in the regression of Step 1 is equal to two. This choice implies that the ATP points of the first three ranked players are investigated in order to define the *strong* players, periods, and eras. Similar results are obtained when the number of regressors in the Step 1 regression slightly increases. Overall, we identify thirty-six *strong* periods, which correspond to five *strong* eras, as illustrated in Table 2.

Several key points emerge. Firstly, we observe many *strong* players during the first and last *strong* eras because these players were alternating in occupying the top three positions of the ATP ranking. Secondly, from August 2019 to December 2023, tennis was in a *weak* era characterized by the absence of *strong* players (as defined in Definition 1). Thirdly, the last *strong* era ended in July 2019 and lasted approximately 133 weeks (roughly 3 years). During the last *strong* era, a total of ten players were part of the group dominating all others. Lastly, Roger Federer is the player with the highest number of weeks among the *strong* players during *strong* eras.

Table 2 The *strong* eras with the relative players (in alphabetic order)

	Start date	End date	Weeks	Players
1	1997-01-13	2000-07-03	181 weeks	A. Agassi, M. Chang, A. Corretja, G. Ivanisevic, Y. Kafelnikov, P. Korda, G. Kuerten, C. Moya, T. Muster, M. Norman, P. Rafter, M. Rios, P. Sampras
2	2001-03-19	2004-03-29	158 weeks	A. Agassi, **R. Federer**, J.C. Ferrero, T. Haas, L. Hewitt, Y. Kafelnikov, G. Kuerten, A. Roddick, M. Safin
3	2007-08-06	2009-05-04	91 weeks	N. Djokovic, **R. Federer**, R. Nadal, A. Roddick
4	2012-07-16	2016-10-03	220 weeks	N. Djokovic, **R. Federer**, D. Ferrer, A. Murray R. Nadal, S. Wawrinka
5	2017-01-09	2019-07-29	133 weeks	M. Cilic, G. Dimitrov, N. Djokovic, **R. Federer**, A. Murray, R. Nadal, J.M. del Potro, M. Raonic, S. Wawrinka, A. Zverev

The player in bold is the player belonging to the *strong* eras for the longest time

3.2 Testing the Hypotheses 1 and 2

In the second part of our analysis, to validate Hypotheses 1 and 2, we examine the relationship between match duration and the occurrence of *strong* players, periods, and eras. This analysis involves several OLS regressions where match duration, expressed in minutes, is the dependent variable.

The independent variables are: *Match between Strong and Not-Strong Players*, *Match between Strong Players*, *ATP Points Difference*, and *Total Games*, *Grass*, and *Hard*. These variables are described as follows:

- ***Match between Strong and Not-Strong Players***: Binary variable equal to one if the match *i* is played during a *strong* period between a *strong* and not-*strong* players. This variable is needed for testing Hypothesis 1. In the regression tables, the variable *Match between Strong and Not-Strong Players* is abbreviated to *M. btw Str./Not-Str. Plyrs*.
- ***Match between Strong Players***: Binary variable equal to one if the match *i* is played among two *strong* players at the time of the match *i*. This variable is needed for testing Hypothesis 2. In the regression tables, the variable *Match between Strong Players* is abbreviated to *M. btw Str. Plyrs*.
- ***ATP Points Difference***: Variable based on the ATP points difference between the two players. The variable uses the following logarithmic transformation:

$$P_{i,t,k} = \log\left(y_{i,t,k}\right),$$

  where $y_{i,t,k}$ is the ATP points of the player in position k for the match i of week t. The variable *ATP Points Difference* is obtained by taking the difference between the transformed points $P_{i,t,k}$ of the players with the largest and smallest amount of points. In the regression tables, the variable *ATP Points Difference* is abbreviated to *ATP Points Diff.*
- ***Total Games***: Variable equal to the total number of games played in the match *i*.
- ***Grass***: Binary variable equal to one if the match *i* is played on grass, and zero otherwise.
- ***Hard***: Binary variable equal to one if the match *i* is played on hard, and zero otherwise.

The OLS regressions' estimates are shown in Table 3. In Column (1), the regression includes only the constant and the binary variable *Match between Strong and Not-Strong Players*. In this first regression, only the constant is significant. In Column (2), the regression uses the constant and the binary variable *Match between Strong Players*. In this case, the coefficient is significant, and, more importantly, the sign is in line with Hypothesis 2: when two *strong* players at the time of the match meet, the match duration, on average, increases by 14.103 minutes, while if the two players are not *strong*, the average duration of the match is 105.092 minutes. Always looking at Table 3, in Column (3), we jointly consider both the *Match between Strong and Not-Strong Players* and *Match between Strong Players* variables. The coefficient of the variable *Match between Strong Players* continues

Table 3 OLS regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Constant	105.128***	105.092***	105.128***	106.863***	−0.799***	1.648***
	(0.138)	(0.132)	(0.138)	(0.201)	(0.198)	(0.203)
M. btw Str./Not-Str. Plyrs	0.656		−0.482	0.582	−1.959***	−1.919***
	(0.484)		(0.503)	(0.51)	(0.223)	(0.215)
M. btw Str. Plyrs		14.103***	14.549***	13.115***	6.371***	6.192***
		(1.664)	(1.728)	(1.73)	(0.755)	(0.731)
ATP Points Diff.				−2.008***	−0.662***	−0.415***
				(0.169)	(0.074)	(0.072)
Total Games					4.236***	4.319***
					(0.007)	(0.007)
Grass						−15.16***
						(0.202)
Hard						−5.721***
						(0.122)
Adjusted R^2	0.000	0.001	0.001	0.002	0.81	0.822
BIC	7.321	7.32	7.32	7.319	5.661	5.594

Dependent variable: *Match Duration* (in minutes)

M. btw Str./Not-Str. Plyrs and *M. btw Str. Plyrs* denote, respectively, the dummy variables *Match between Strong and Not-Strong Players* and *Match between Strong Players*

Sample period: January 1990 to November 2023. Number of matches: 86,823

Standard errors are in parentheses

*, ** and *** represent the significance at levels 10%, 5%, 1%, respectively

to be significant and coherent with Hypothesis 2. Column (4) adds the variable *ATP Points Difference* to the regressors in Column (3). Interestingly, the sign of the coefficient associated with *ATP Points Difference* is negative, as expected: when the difference between the ATP points of the players increases, the match duration, on average and controlling for all the other variables, decreases. In Column (5), the set of regressors includes the variable *Total Games*, which is significant and with a positive sign: a higher number of games implies a longer match duration, on average, and controlling for all the other variables. Furthermore, in Column (5), the signs of the variables *Match between Strong and Not-Strong Players* and *Match between Strong Players* are, respectively, negative and positive, and all the coefficients are significant. In Column (6), dummy variables related to the playing surface are included, with the omitted variable being *Clay*. Remarkably, all coefficients remain significant, and the signs remain unchanged with respect to the coefficients in Column (5). When a *strong* player, at the time of the match, faces a not-*strong* player, the match duration decreases by approximately 1.9 minutes, on average, and controlling for all the other variables. When two *strong* players at the time of the match face, the match duration increases by 6.192 minutes, on average, and keeping all the other variables constant. Playing on grass reduces the match duration by 15.16 minutes, on average, and keeping all the other variables constant. The last two rows of Table 3 present some diagnostic measures. The adjusted R^2 of the model in Column (6) is the highest at 0.822, indicating the strongest explanatory power. Additionally, the BIC suggests that the model in the last column is the most suitable.

In conclusion, we found strong evidence in favor of our Hypotheses 1 and 2: during *strong* periods, when a *strong* player faces a not-*strong* player, the match duration tends to decrease, while when two *strong* players (at the time of the match) meet, the match duration increases. In the regressions in Columns (5) and (6) of Table 3, the coefficients used to evaluate the Hypotheses 1 and 2 are always significant, even when including the other controlling variables in the regressions.

3.3 Robustness check

In this paragraph, we first verify the robustness of our results related to the *strong* eras by changing the value of T_{in} parameter. We consider a number of alternative choices of T_{in}, that is from $T_{in} = 84$ and $T_{in} = 96$ (the original choice was $T_{in} = 90$). The results are shown in Fig. 2. Remarkably, the length and the number of *strong* eras are unchanged if compared to the original choice of $T_{in} = 90$.

The robustness of our results concerning the Hypotheses 1 and 2 is also further investigated through regressions involving matches played exclusively in Grand Slams (Panel A of Table 4) and in ATP Tour Masters 1000 (Panel B of Table 4). Interestingly, all the previous results in Columns (5) and (6) of Table 3 are confirmed. It is worth noting that, restricting the analysis to Grand Slams and ATP Tour Masters 1000, when usually almost all top-ten players participate, even the regressions in Column (1) present significant coefficients for the variable *Match between Strong and Not-Strong Players*. This happens because, in Grand Slams and ATP Tour Masters 1000, there are more matches involving *strong* and not-*strong* players

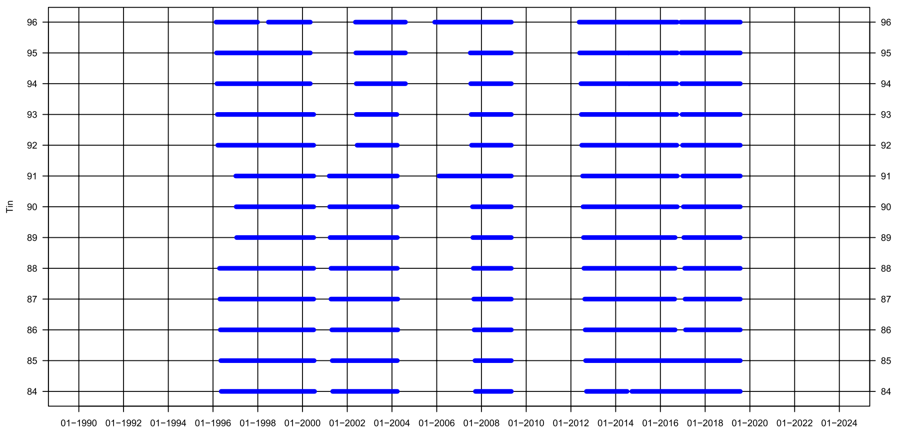

Fig. 2 Robustness check for eras. The figure illustrates the identified strong eras (blue segments) by different choices of the T_{in} parameter

Table 4 Robustness check for OLS estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Grand Slam matches						
Constant	149.127***	148.089***	149.127***	156.456***	−5.766***	−0.961
	(0.417)	(0.394)	(0.416)	(0.633)	(0.843)	(0.854)
M. btw Str./Not-Str. Plyrs	−7.155***		−9.642***	−4.602***	−0.676	−1.11*
	(1.22)		(1.267)	(1.3)	(0.647)	(0.619)
M. btw Str. Plyrs		20.235***	28.838***	22.05***	10.936***	9.916***
		(3.925)	(4.078)	(4.069)	(2.024)	(1.939)
ATP Points Diff.				−7.6***	−0.092	0.154
				(0.497)	(0.25)	(0.24)
Total Games					4.255***	4.305***
					(0.021)	(0.02)
Grass						−17.687***
						(0.52)
Hard						−4.787***
						(0.451)
Obs.	14,121	14,121	14,121	14,121	14,121	14,121
Adjusted R^2	0.002	0.002	0.006	0.022	0.758	0.778
BIC	7.681	7.681	7.677	7.661	6.263	6.177
Panel B: Master 1000 matches						
Constant	99.504***	99.076***	99.504***	101.221***	−8.176***	−4.647***
	(0.257)	(0.241)	(0.257)	(0.379)	(0.474)	(0.483)
M. btw Str./Not-Str. Plyrs	−2.651***		−3.529***	−2.534***	−1.991***	−2.254***
	(0.703)		(0.736)	(0.753)	(0.35)	(0.343)
M. btw Str. Plyrs		5.598***	8.699***	7.43***	6.146***	6.074***
		(2.075)	(2.172)	(2.18)	(1.014)	(0.994)
ATP Points Diff.				−2.134***	0.822***	0.939***
				(0.347)	(0.162)	(0.159)
Total Games					4.62***	4.647***
					(0.019)	(0.018)
Hard						−6.187***
						(0.234)
Obs.	17,101	17,101	17,101	17,101	17,101	17,101
Adjusted R^2	0.001	0	0.002	0.004	0.785	0.793
BIC	6.884	6.885	6.884	6.881	5.35	5.31

Dependent variable: *Match Duration* (in minutes)

M. btw Str./Not-Str. Plyrs and *M. btw Str. Plyrs* denote, respectively, the dummy variables *Match between Strong and Not-Strong Players* and *Match between Strong Players*

Sample period: January 1990 to November 2023

Standard errors are in parentheses

*, ** and *** represent the significance at levels 10%, 5%, 1%, respectively

with respect to matches analyzed in Table 3. In particular, the average duration of a match involving two not-*strong* players is 149.127 and 99.504 minutes, respectively, for Grand Slam and ATP Tour Masters 1000 matches (that is, when the dummy variable *M. btw Str./Not-Str. Plyrs* in Column (1) is zero). When the match involves a *strong* and a not-*strong* players (hence, the dummy variable *M. btw Str./Not-Str. Plyrs* in Column (1) is one), the average duration decreases by 7.155 and 2.651 minutes for Grand Slam and ATP Tour Masters 1000 matches, respectively. Independently of the type of tournament, the signs of the *Match between Strong and Not-Strong Players* and *Match between Strong Players* variables are always negative and positive, respectively, as expected. Moreover, almost all the other control variables remain significant and, in most cases, with their expected signs.

To conclude, we find strong evidence in favor of the proposed Hypotheses 1 and 2: when a *strong* player faces a not-*strong* competitor, the match duration, on average, decreases, while when two *strong* players meet, the match duration generally increases.

4 Discussion and conclusion

In this work, we answered methodologically and empirically the questions “Is it possible to find *strong* players who dominate all the others for a given amount of time?”, “How many *strong* periods have been in tennis?”, and “Which is the relationship between match duration and the existence of *strong* players, periods, and eras?”. By employing cointegration analysis of ATP points among the top players for the first time in this context, covering the period from 1990 to 2023, we identified several *strong* players in the detected thirty-six *strong* periods, corresponding to five *strong* eras. The rationale behind our methodological approach is as follows. The ATP points of the top ten players are not stationary time series, but some of these may be cointegrated for a certain period. Applying the Engle–Granger two-step cointegration analysis dynamically to the ATP points of the top-ranked players, we found that some ATP points time series are cointegrated during specific time frames. This means that the ATP points under investigation share a common trend. Therefore, the players holding the respective ranking positions (of the cointegrated ATP points) are labeled as *strong*, in the sense that, by sharing a common trend relative to those in lower positions, they consistently and jointly earn points and dominate over the players in lower positions. Notably, the last detected *strong* era ended in July 2019, two weeks after the epic Wimbledon 2019 Final between Roger Federer and Novak Djokovic. Interestingly, the player with the largest participation (that is, 602 weeks) in the *strong* eras is Roger Federer. The so-called ‘Big Three’, listed alphabetically, Novak Djokovic, Roger Federer, and Rafael Nadal, were prominent in the last three *strong* eras. Interestingly, according to our definitions and findings, tennis, up to December 2023, is experiencing a *weak* era, possibly related to Federer’s retirement and Nadal’s prolonged injuries.

In view of the research conducted by Baker and McHale (2014) and Baker and McHale (2017), the approach proposed for detecting *strong* players, periods, and eras represents a generalization of methods for identifying the greatest players of all time. In this regard, the achievements and records of players should also consider the strength of their opponents at that time. Winning most of twenty Grand Slam titles, as Federer did during the identified *strong* eras, is much more challenging than winning slams during *weak* eras.

Another key point investigated in this work is recognizing statistically-founded rivalries. In the five *strong* eras detected, there is statistical evidence to state that not only the well-known Federer–Nadal rivalry is noteworthy, but also other rivalries like Federer–Djokovic, Djokovic–Nadal are present, among others. Identifying statistically-founded rivalries among *strong* players in a *strong* era is of interest for several reasons. Firstly, in the tennis-related literature, the works of Chmait et al. (2020) and Konjer et al. (2017) have already found that matches between top players lead to increased ticket sales and television audiences. The attraction of the audience for the matches between *strong* players is therefore evident and also surged by the resulting match uncertainty, which increases demand for sports (Borland and MacDonald 2003; Simmons 2006; Forrest and Simmons 2002, among others). On the other hand, the occurrence of a *weak* era is likely to reduce the attractiveness of tennis. Secondly, from an economic standpoint, the anticipation of matches between two *strong* players (in a *strong* era) can significantly enhance global visibility and attract increased sponsorship. Moreover, the presence of *strong* players in a *strong* era could stimulate the adoption of *ad-hoc* marketing opportunities (Ambrose and Schnitzlein 2017), like the signature logo created by Nike for Federer and Nadal in the mid-2000s, where these two players were dominating.

While *strong* player rivalries and *strong* eras in tennis are widely acknowledged for their significant marketing appeal, understanding and identifying *weak* eras also play a crucial role in strategic decision-making for sports organizations. Recognizing periods characterized by lower competitive intensity can provide organizations with opportunities to negotiate advantageous deals with their most valuable players. During these phases, where the overall competitive landscape may be less intense, organizations can capitalize on the unique market positioning and commercial potential of their top athletes. By leveraging statistical insights into *weak* eras, organizations can tailor sponsorship agreements, endorsement contracts, and promotional campaigns to maximize the marketability and visibility of athletes, who, in the future, may become *strong* players. This strategic approach not only strengthens partnerships between players and organizations but also enhances the overall brand positioning and profitability in the competitive sports market.

Finally, the third key aspect we contributed to was the relationship between the match duration and *strong* players/periods/eras. We found that the occurrence of a match between a *strong* and not-*strong* players decreases the match duration, on average. While, if two *strong* players meet, the match duration increases, on average. These findings are not only robust to the inclusion of several control variables in the OLS regressions but also to the analysis repeated exclusively considering matches played in Grand Slams or ATP Tour Masters 1000. Therefore, the relationship between the match duration and *strong* players/periods/eras could help tournament directors in defining appropriate scheduling for the most appealing and eventually uncertain events.

For future research, the methods used to identify *strong/weak* periods/eras in this study could also be applied to the female tennis circuit. Furthermore, the approach here may be applicable in various sports beyond tennis. Last but not least, the relationship between the *strong* periods/eras and match uncertainty (recently investigated by Özaydın and Könecke (2024), among others) could be assessed.

Acknowledgments We are grateful to Philippe Vieu, Pierpaolo D'Urso, Michele Gallo, Paola Zuccolotto, and the two anonymous referees for their constructive comments, which greatly helped us improve our contribution. We also thank the participants of IES 2023 — Statistical Methods for Evaluation and Quality: Techniques, Technologies, and Trends, held in Pescara, Italy, for their valuable feedback.

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716–723

Akhanli SE, Hennig C (2023) Clustering of football players based on performance data and aggregated clustering validity indexes. Int J Perform Anal Sport 19:103–123

Albert J, Glickman ME, Swartz TB, Koning RH (2017) Handbook of statistical methods and analyses in sports. Chapman and Hall-CRC, Boca Raton

Alfano V, Cicatiello L, Gaeta GL, Gallo M, Rotondo F (2021) Three is a magic number: evidence on the effects of the application of the three-point rule in Italy's Serie A. J Sports Econ 22(3):329–356

Ambrose SC, Schnitzlein N (2017) What makes for the best rivalries in individual sports and how can marketers capitalize on them? Sport Mark Quart 26(4):223–234

Ang JB (2007) CO2 emissions, energy consumption, and output in France. Energy Policy 35(10):4772–4778

Angelini G, Candila V, De Angelis L (2022) Weighted Elo rating for tennis match predictions. Eur J Oper Res 297(1):120–132

Arcagni A, Candila V, Grassi R (2023) A new model for predicting the winner in tennis matches based on the eigenvector centrality. Ann Oper Res 325(1):615–632

Bai L, Gedik R, Egilmez G (2023) What does it take to win or lose a soccer game? A machine learning approach to understand the impact of game and team statistics. J Oper Res Soc 74:1690–1711

Baker RD, McHale IG (2013) Forecasting exact scores in National Football League games. Int J Forecast 29(1):122–130

Baker RD, McHale IG (2014) A dynamic paired comparisons model: who is the greatest tennis player? Eur J Oper Res 236(2):677–684

Baker RD, McHale IG (2017) An empirical Bayes model for time-varying paired comparisons ratings: who is the greatest women's tennis player? Eur J Oper Res 258(1):328–333

Baumer BS, Matthews GJ, Nguyen Q (2023) Big ideas in sports analytics and statistical tools for their investigation. Computat Stat 15:e1612

Berry SM, Reese CS, Larkey PD (1999) Bridging different eras in sports. J Am Stat Assoc 94:661–686

Borland J, MacDonald R (2003) Demand for sport. Oxf Rev Econ Policy 19(4):478–502

Box GE, Jenkins GM, Reinsel GC, Ljung GM (2015) Time series analysis: forecasting and control, 5th edn. Wiley, Hoboken

Breznik K (2015) Revealing the best doubles teams and players in tennis history. Int J Perform Anal Sport 15(3):1213–1226

Candila V, Palazzo L (2020) Neural networks and betting strategies for tennis. Risks 8(3):68

Cefis M, Carpita M (2022) The higher-order PLS-SEM confirmatory approach for composite indicators of football performance quality. Comput Stat. https://doi.org/10.1007/s00180-022-01295-4

Chmait N, Robertson S, Westerbeek H, Eime R, Sellitto C, Reid M (2020) Tennis superstars: the relationship between star status and demand for tickets. Sport Manag Rev 23(2):330–347

Cohen-Zada D, Krumer A, Shapir OM (2018) Testing the effect of serve order in tennis tiebreak. J Econ Behav Organ 146:106–115

de Paula Oliveira T, Newell J (2024) A hierarchical approach for evaluating athlete performance with an application in elite basketball. Sci Rep 14:1717

Del Corral J, Prieto-Rodríguez J (2010) Are differences in ranks good predictors for Grand Slam tennis matches? Int J Forecast 26(3):551–563

Dickey DA, Fuller WA (1979) Distribution of the estimators for autoregressive time series with a unit root. J Am Stat Assoc 74(366a):427–431

Dobson S, Goddard JA (1996) The demand for football in the regions of England and Wales. Reg Stud 30(5):443–453

Dona NE, Gill PS, Swartz TB (2024) What does rally length tell us about player characteristics in tennis? J R Stat Soc Ser A: Stat Soc. https://doi.org/10.1093/jrsssa/qnae027

D'Urso P, De Giovanni L, Vitale V (2023) A robust method for clustering football players with mixed attributes. Ann Oper Res 325(1):9–36

Elo AE (1978) The rating of chessplayers, past and present. ARCO Pub, New York

Engle RF, Granger CW (1987) Co-integration and error correction: representation, estimation, and testing. Econometrica 55:251–276

Forrest D, Simmons R (2002) Outcome uncertainty and attendance demand in sport: the case of English soccer. J R Stat Soc Ser D: Stat 51(2):229–241

Gabrio A (2021) Bayesian hierarchical models for the prediction of volleyball results. J Appl Stat 48(2):301–321

George G, Panagiotis Z (2008) Statistical analysis of men's FIVB beach volleyball team performance. Int J Perform Anal Sport 8(1):31–43

Gorgi P, Koopman SJS, Lit R (2019) The analysis and forecasting of tennis matches by using a high dimensional dynamic model. J R Stat Soc Ser A Stat Soc 182(4):1393–1409

Hamilton JD (1994) Time series analysis. Princeton University Press, Princeton

Hansen BE (1992) Efficient estimation and testing of cointegrating vectors in the presence of deterministic trends. J Econom 53(1–3):87–121

Harris DJ, Vine SJ, Eysenck MW, Wilson MR (2021) Psychological pressure and compounded errors during elite-level tennis. Psychol Sport Exerc 56:101987

Khan H, Toh RS, Chua L (2005) Tourism and trade: cointegration and granger causality tests. J Travel Res 44(2):171–176

Kirschstein T, Liebscher S (2019) Assessing the market values of soccer players—a robust analysis of data from German 1. and 2. Bundesliga. J Appl Stat 46:1336–1349

Konjer M, Meier HE, Wedeking K (2017) Consumer demand for telecasts of tennis matches in Germany. J Sports Econ 18(4):351–375

Kovalchik S (2020) Extension of the Elo rating system to margin of victory. Int J Forecast 36:1329–1341

Lisi F, Zanella G (2017) Tennis betting: can statistics beat bookmakers? Electron J Appl Stat Anal 10(3):790–808

López-Serrano C, Zakynthinaki MS, Mon D, Molina-Martin JJ (2024) Introducing the technical individual contribution coefficient: a metric for evaluating performance in elite volleyball. Int J Perform Anal Sport 24:204–217

McHale I, Morton A (2011) A Bradley–Terry type model for forecasting tennis match results. Int J Forecast 27(2):619–630

Milekhina A, Breznik K, Restaino M (2023) Gender comparison of in-match psychological traits of tennis players: dynamic network analysis. Statistica Applicata. https://doi.org/10.26398/IJAS.0035-011

Morgulev E, Azar OH, Lidor R (2018) Sports analytics and the big-data era. Int J Data Sci Anal 5:213–222

Musa M (2023a) Tennis viewership statistics. https://tennisracketball.com/guide/how-many-people-play-tennis/. Accessed 25 Jan 2024

Musa M (2023b) Tennis viewership statistics. https://tennisracketball.com/guide/tennis-viewership-statistics/. Accessed 25 Jan 2024

Özaydın S, Könecke T (2024) Match-level uncertainty in professional tennis revisited—a novel approach applied for the time between 2010 and 2019. J Sports Econ. https://doi.org/10.1177/15270025241233549

Peiris H, Dona NE, Swartz T (2024) Analysis of the impact of unforced errors in tennis arXiv:2407.19321

Phillips PC, Ouliaris S (1990) Asymptotic properties of residual based tests for cointegration. Econometrica 58:165–193

Radicchi F (2011) Who is the best player ever? A complex network analysis of the history of professional tennis. PLoS ONE 6(2):e17249

Sandri M, Zuccolotto P, Manisera M (2020) Markov switching modelling of shooting performance variability and teammate interactions in basketball. J R Stat Soc: Ser C: Appl Stat 69(5):1337–1356
Schumaker RP, Solieman OK, Chen H (2010) Sports data mining. Springer, New York
Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6(2):461–464
Simmons R (2006) Chapter 8: the demand for spectator sports. Edward Elgar Publishing, Cheltenham
Tea P, Swartz TB (2023) The analysis of serve decisions in tennis using Bayesian hierarchical models. Ann Oper Res 325:633–648
Wang Q, Yao N (2024) Light imaging detection based on cluster analysis for the prevention of sports injury in tennis players. Opt Quant Electron 56:191
Wilkens S (2021) Sports prediction and betting models in the machine learning age: the case of tennis. J Sports Anal 7(2):99–117
Zuccolotto P, Sandri M, Manisera M (2023) Spatial performance analysis in basketball with CART, random forest and extremely randomized trees. Ann Oper Res 325(1):495–519

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Kristijan Breznik[1,2] · Vincenzo Candila[3] · Antonina Milekhina[4] · Marialuisa Restaino[3]

✉ Vincenzo Candila
vcandila@unisa.it

Kristijan Breznik
kristijan.breznik@mfdps.si

Antonina Milekhina
antonina.milekhina@yandex.ru

Marialuisa Restaino
mlrestaino@unisa.it

1 International School for Social and Business Studies, Mariborska cesta 7, 3000 Celje, Slovenia

2 Faculty of Environmental Protection, Trg mladosti 7, 3320 Velenje, Slovenia

3 Department of Economics and Statistics, University of Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano (Salerno), Italy

4 Independent researcher, Moscow 610101, Russia